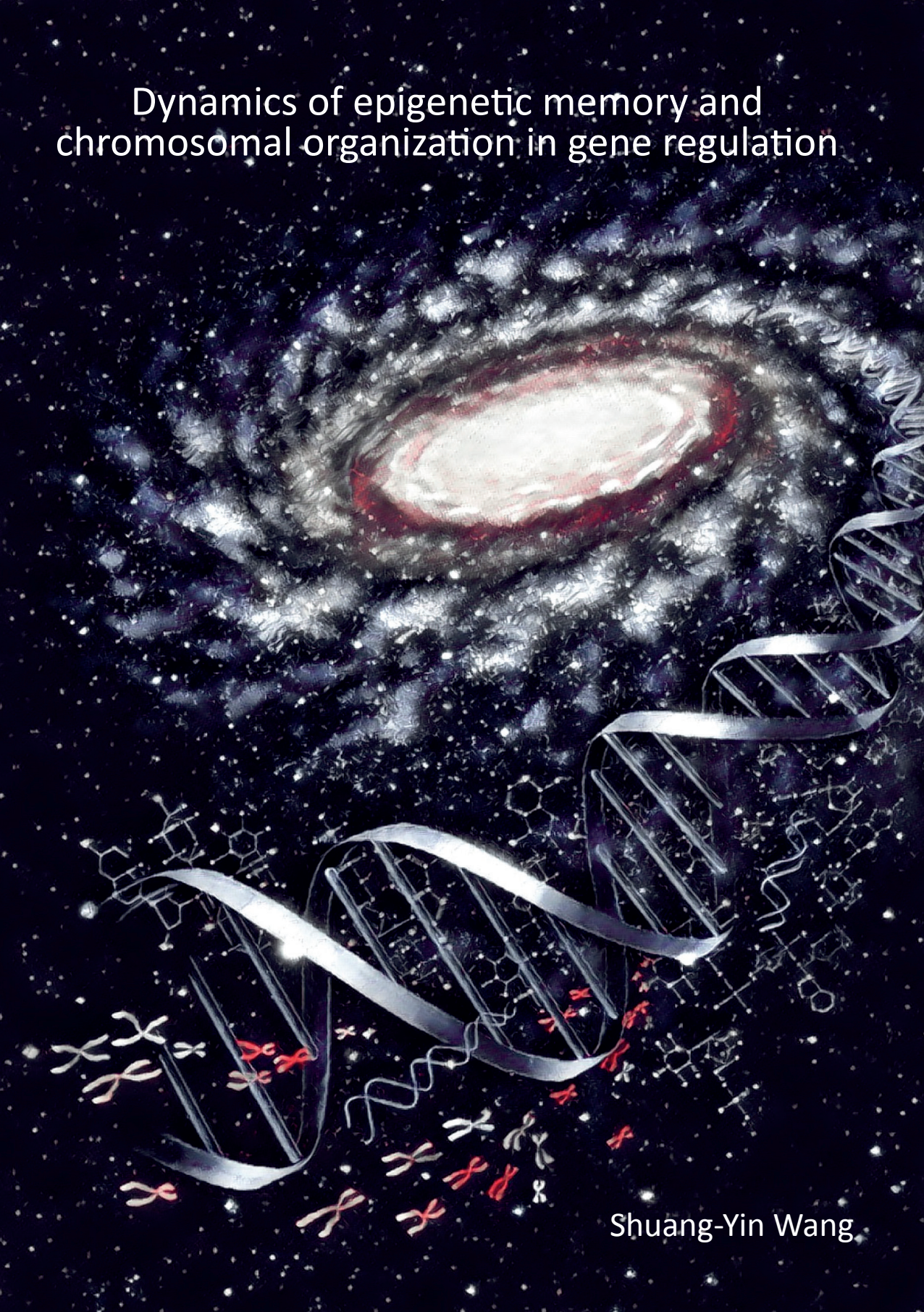
Dynamics of epigenetic memory and chromosomal organization in gene regulation



Shuang-Yin Wang

Dynamics of epigenetic memory and chromosomal organization in gene regulation

Shuang-Yin Wang

The research presented in this thesis was performed at the Department of Molecular Biology, Radboud Institute for Molecular Life Sciences, Radboud University Nijmegen, The Netherlands and was supported by the 'Nederlandse Organisatie voor Wetenschappelijk Onderzoek' (NWO), European Research Council (ERC), and European Union's Seventh Framework Programme (FP7). The computational work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative.

ISBN: 978-94-92380-62-3
Design cover: Shuang-Yin Wang
Design inside: ProefschriftOntwerp.nl
Printed by: Ipskamp

Dynamics of epigenetic memory and chromosomal organization in gene regulation

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 11 september 2017
om 10.30 uur precies

door

Shuang-Yin Wang
geboren op 17 februari 1986
te Wuhan, P.R. China

Promotor:
Prof. dr. ir. H. G. Stunnenberg

Manuscriptcommissie:
Prof. dr. G. J. C. Veenstra
Prof. dr. F. Grosveld (Erasmus MC)
Prof. dr. J.J. Schuringa (Universiteit Groningen)

Dynamics of epigenetic memory and chromosomal organization in gene regulation

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Monday, September 11, 2017
at 10.30 hours

by

Shuang-Yin Wang
Born on February 17, 1986
in Wuhan, P.R. China

Supervisor:
Prof. dr. ir. H. G. Stunnenberg

Doctoral Thesis Committee:
Prof. dr. G. J. C. Veenstra
Prof. dr. F. Grosveld (Erasmus MC)
Prof. dr. J.J. Schuringa (University of Groningen)

Contents

Chapter 1

Introduction

Deciphering the "language" of the genome is a crucial step to understand life on earth. The Human Genome Project (HGP) initiated the era of Omics, and the rapid method development has greatly improved our knowledge of genome. Epigenomics, one of the active Omics fields, has largely increased our genomic vocabulary of the genomic and functional DNA elements. It has revealed that amongst others the local chromatin configuration at these DNA elements plays a pivotal role in generating cell types and modulating cell functions (Bernstein et al., 2007; Consortium, 2012; Ernst et al., 2011; Goldberg et al., 2007; Kouzarides, 2007; Neph et al., 2012a), and has been linked to health and disease (Arrowsmith et al., 2012; Calvanese et al., 2009; Egger et al., 2004; Feil and Fraga, 2012; Kelly et al., 2010; Portela and Esteller, 2010). Most recently, studies of the three-dimensional (3D) genome, which can be analogized to genomic syntax, has progressively drawn our attention to chromatin physical communication and begun yielding exciting insights into the hierarchical chromatin organization and its contribution to gene regulation by facilitating the contacts between separated genomic regions (Dekker et al., 2013; Denker and Laat, 2016; Dixon et al., 2012; Lieberman-Aiden et al., 2009). This chapter introduces the general conceptions, existing techniques and the aims of our studies.

General conceptions

How a genome governs phenotypic plasticity, has long been the subject of research.

The genome that encodes the genetic blueprint of an organism consists of linear polymers comprising a colossal amount of deoxyribonucleotides (ribonucleotides in some viruses) ranging from thousands in viruses to billions in eukaryotes.

It is extremely fascinating and mind boggling that multicellular eukaryotes, such as the metazoans, are composed of a multitude of different cell types that have virtually the same genomic information yet have distinct morphologies and functions. In addition, the cells of a metazoan and for that matter unicellular organisms have the ability to respond to cues from the environment. Thus, differentiation and other cellular responses must be carried out and maintained by a precisely spatiotemporal regulation of genes, yet remain plastic to respond to a changing environment. Genome sequencing has revealed that, in most eukaryotic genomes, only very a small part of the genome (~2% in human and mouse) is used to encode proteins (Goldberg et al., 2007; Harrow et al., 2012; Venter et al., 2001). This raises the question what the rest is used for. Originally, a major part of the genome was thought to be junk. It has become clear however, that non-coding genomic regions likely have a function in regulation of the transcriptional activities of genes. The non-coding part of the genome, like the coding part, is modified with specific chemical tags that mark different regions with specific functions and instruct the cellular machinery to execute a particular

cellular function. Such modifications define so called cis-regulatory DNA elements (CREs). A variety of techniques have been developed to identify the distinctly marked CREs and assess the function of these genomic DNA elements. CREs are often found at a large distance from the genes and yet appear to regulate gene expression. This begs the question how such distant elements communicate with the transcriptional machinery located at the start site of the gene, the promoter. The development of chromosome conformation capture (3C) and genome-wide derivatives enable to link CREs to their target regions, and unveiled a complex hierarchical compaction and organization of eukaryotic genomes (Table 1). Epigenomics and 3D genomics opened the door to study how phenotypes are created and sustained in populations over time. Yet, due to the high diversities of cell types and the requirement of abundant starting material, in depth genome-wide study have mainly been performed in depth on cell lines and a handful of primary cell types from mammals over the last few years (Bernstein et al., 2005; Consortium, 2012; Marks et al., 2012; Roadmap Epigenomics Consortium et al., 2015; Saeed et al., 2014; Stunnenberg et al., 2016).

Epigenome — primary structure of chromosome

Until the discovery of the molecular basis of epigenomes, the inheritable information which defines different cell types remained a puzzle long after the establishment of primary structure of DNA, the double-helix model, in the early 1950's. Epigenome refers to the functionally relevant changes to the genome that do not alter the nucleotide sequences but are inheritable upon cell division (Bernstein et al., 2007; Margueron and Reinberg, 2010). In eukaryotic cells, the nucleosome which provides the first layer of DNA compaction is the basic unit of chromatin consisting of a segment of DNA wrapped around a histone octamer containing two copies each of the core histones H2A, H2B, H3, and H4 (Khorasanizadeh, 2004). Both the DNA and histone proteins are decorated with various reversible chemical modifications to set up a localized chromatin configuration which has a major influence on transcriptional activities of genes (Bannister and Kouzarides, 2011). Taken together, the double-helix DNA of genome provides an elegant way to accurately transfer genetic information to the next generation, whereas the reversible and inheritable chemical modifications to chromatin add the ability to construct diverse gene expression programs for distinct phenotypes and functions.

DNA modification

In nature, DNA is made up from four nucleotides, each known by its own letter—A, G, C, and T. Several DNA nucleotide modifications are often found on the usual letters in the DNA sequences (Plongthongkum et al., 2014). Generally, these modified nucleotides are not considered as usual nucleotides, because the DNA replication machinery itself could not

preserve those variants in the daughter strands after every cellular DNA replication cycle (Chen and Riggs, 2011; Smith and Meissner, 2013).

DNA methylation is the most common DNA modification by which methyl groups are added to nucleotides. In mammals, methylation at the 5th carbon position of cytosine is the most abundant DNA modification (Suzuki and Bird, 2008). Most 5mCs are found in the context of CpG dinucleotides, where the C residues on the two strands of this short palindromic sequence are usually both methylated. 5mC can be found throughout vertebrate genomes. Its patterns are stable during replication and discriminate different cell types. The DNA methyltransferase DNMT1 is responsible for maintenance of DNA methylation patterns following DNA replication, whereas DNMT3A and DNMT3B are involved in deposition at a new site (Wu and Zhang, 2014).
DNA methylation often acts like switches of genes controlling which genes are turned on and off (Suzuki and Bird, 2008). It mainly affects gene activities by its interplays with other proteins. Methylated DNA can be recognized and bound by proteins known as methyl-CpG-binding domain proteins (MBDs) which then recruit additional proteins such as the polycomb group proteins to create an inactive chromatin configuration for gene repression (Zhu et al., 2016). Alternatively, but rarely, the DNA methylation itself may physically impede the binding of certain transcription factors (TFs) (Choy et al., 2010).

It has been widely accepted that 5-methylcytosine (5mC) is essential for normal development and is associated with a number of key processes including genomic imprinting (Bartolomei and Ferguson-Smith, 2011), X-chromosome inactivation (Sharp et al., 2011) and repression of repetitive elements (Habibi et al., 2013; Smith et al., 2014). Hence, abnormal changes in DNA methylation patterns may link to human diseases (Bernstein et al., 2007; Heyn and Esteller, 2012). Using bisulfite sequencing and other techniques, a number of genomic regions have been associated with cancer and aging showing local hypomethylation and hypermethylation at specific genes (Berman et al., 2012; Bernstein et al., 2007; Jones et al., 2016; Varley et al., 2013). Although the precise mechanism is not completely understood, these regions such as differential methylated regions (DMRs) can be potentially used as biomarkers for the diagnosis of disease progression. Some other DNA modifications such as 5-hydroxymethylcytosine (5hmC) are also found in vertebrate genomes (Ito et al., 2011; Tahiliani et al., 2009; Whalen et al., 2016; Yu et al., 2012). However, their roles in gene regulation have not been extensively explored.

DNA accessibility

While nucleosomes help in DNA compaction, they also provide protection to DNA from access. To utilize the information stored in the genome, DNA sequences have to be accessible by other molecules residing in the nucleus such as transcription factors (TFs). To systematically measure DNA accessibility, several techniques have been developed based on the idea that DNA at nucleosome free regions can be cleaved or modified by enzymes such as DNase I and transposase Tn5 (Meyer and Liu, 2014). Despite certain observed biases, these techniques produced strong cross-validation showing 1~2% of the genome as accessible regions (often called open chromatin regions) in individual cell types (He et al., 2014).

Exposed DNA has long been believed to indicate activated regulatory regions such as promoters and enhancers. Hence, DNA accessibility assay provides a highly effective way to study transcriptional regulation at genome wide level. Indeed, it has been shown by ChIP–seq that transcription factor binding displays high correlation with DNA accessibility (Neph et al., 2012a; Thurman et al., 2012; Whalen et al., 2016). Open chromatin regions likely are created by a special class of transcription factors called pioneer factors. These factors bind to nucleosomal DNA and recruit additional factors to evict the nucleosome, thereby making the DNA accessible to other transcription factors (Bell et al., 2011; Struhl and Segal, 2013; Zaret and Carroll, 2011). Alternatively, some transcription factors might directly open DNA by competing with histone octamer for DNA binding during spontaneous unwrapping and rebinding of the nucleosome such as RNA polymerase II (Li et al., 2005). Either way, DNA accessibility assay captures these open chromatin regions, and therefore provides a robust way to detect regulatory elements independently of any given transcription factor.

DNA accessibility assay such as DNaseI-seq and ATAC-seq also enabled a genome-wide approach to identify DNA footprints (Buenrostro et al., 2013; Neph et al., 2012a, 2012b). This phenomenon occurs because binding of a protein to DNA often protects that DNA from enzymatic cleavage. A DNA footprint can be observed as a small depletion — typically only a few base pairs long — within a DNA accessible region. This information can be used to predict which DNA-sequence-specific transcription factors bind to a given open chromatin region as it produces a specific footprint, hence providing the evidence for the interactions between DNA and transcription factors of interests.

Histone modification

Many researchers are fascinated by histone modifications because of their diversity, reversibility and heritability. A histone modification is a covalent post-translational modification (PTM) to histone proteins which includes methylation, phosphorylation, acetylation, ubiquitylation, and sumoylation. Some of the known histone modifications have been suggested to act in diverse biological processes such as gene regulation, DNA repair

and chromosomal condensation (Bannister and Kouzarides, 2011). Most well-studied histone modifications are involved in gene regulation and primarily found on unstructured tails of histone proteins, such as methylation and acetylation.

Histone methylation is defined as the addition of methyl groups to the amino acid residues of histone proteins by histone methyltransferases (HMTs) (Bartke et al., 2010). Many studies concentrated on the histone methylation at histone H3 lysine 4 (H3K4), H3K9, H3K27, H3K36, H3K79 and H4K20 because of their high correlation with gene activities (Black et al., 2012; Martin and Zhang, 2005). Lysines can be monomethylated (me1), dimethylated (me2) or trimethylated (me3). The SET-domain-containing proteins and DOT1-like proteins have been shown to act as "writers" that methylate lysines. Histone methylation, like many other histone modifications, is reversible (Bannister and Kouzarides, 2005; Wu and Zhang, 2014). Two main families of histone demethylases have been identified as "erasers" that demethylate methyl lysines, including flavin adenine dinucleotide (FAD)-dependent amine oxidase, and Fe(II) and α-ketoglutarate-dependent dioxygenase. Mono, di and tri-methyl lysine can be recognized and distinguished by "reader" proteins containing chromo, PHD or some other domains, and potentially confer different biological meanings (Black et al., 2012; Martin and Zhang, 2005).

Histone acetylation occurs through the addition of an acetyl group from acetyl coenzyme A to a lysine residue (Fukuda et al., 2006). Histone acetylation and deacetylation are catalyzed by "histone acetyltransferases" (HATs) or "histone deacetylases" (HDACs), respectively (Marmorstein and Zhou, 2014; Seto and Yoshida, 2014). Histone acetylation removes the positive charge of the histone octamer, thereby decreasing the interaction between the histone proteins and the negatively charged phosphate groups in DNA. This results in a more relaxed chromatin structure that assists the access of diverse DNA-binding proteins to the DNA (Grunstein, 1997; Tessarz and Kouzarides, 2014).

Accumulating genome-wide mapping of histone modifications revealed highly stereotypical patterns with distinct combinations enriched at different classes of regulatory elements (Calo and Wysocka, 2013; Ernst et al., 2011; Zhou et al., 2011). The patterns of methylation and acetylation on lysines are used to annotate chromatin states and to identify enhancers and promoters in a given cell type. Promoters are typically marked by high H3K4me3 and low H3K4me1 whereas for enhancers this is the other way around. Both promoters and enhancers usually are additionally marked by H3K27ac upon activation. By contrast, repressed promoters and enhancers are often occupied by nucleosomes that contain H3K27me3, which has been linked to chromatin repression by polycomb-group proteins. In addition, H3K36me3 particularly marks areas of transcription elongation, whereas H3K9me3 typically marks transcriptionally silenced heterochromatin regions.

Histone variants

A number of variants are also found for the core histones H3, H2A and H2B, as well as for the linker histone H1 (Bartke et al., 2010; Weber and Henikoff, 2014). These variants differ from the canonical histones by the alteration of one or a few amino acids or by the addition of larger domains. Some variants do not introduce any considerable structural alteration to the nucleosomes such as H3.3, whereas variant like H2A.Z influences the stability of H2A.Z-containing nucleosomes (Bartke et al., 2010). Generally, histone variants have a specific abundant pattern, genomic localization and species distribution, and confer novel functional properties on the nucleosome affecting chromatin remodeling and histone modifications.

Enhancer

Enhancers can be analogous to adjectives or adverbs in human languages. Enhances were initially described as short DNA fragments which have the ability to "modify" the expression of target genes (Banerji et al., 1981). The hallmark of enhancers is their functional independence of genomic distance and orientation relative to the promoters of target genes (Banerji et al., 1981; Blackwood and Kadonaga, 1998; Plank and Dean, 2014). Further studies demonstrated that enhancers locate in regions containing short DNA motifs that act as binding sites for DNA-sequence-specific transcription factors (TFs) (Mitchell and Tjian, 1989). These TFs can recruit co-activators or co-repressors that then jointly determine the activity of an enhancer. As a result, enhancer regions usually coincide with open chromatin regions which can be identified by DNA accessibility assay (Consortium, 2012; Plank and Dean, 2014; Roadmap Epigenomics Consortium et al., 2015; Yao et al., 2015).

Understanding enhancers has become an area of great interest, as increasing evidence has continuously strengthened their roles in regulating gene expression not only in differentiation and development but also in evolution and disease. Enhancers mediate both spatial and temporal control of gene activities and hence display cell-type specific repertoires. To identify enhancers, the coactivator P300, which is a histone acetyltransferase often binding at enhancer regions, was initially used as an indicator (Calo and Wysocka, 2013). Recently, increasing knowledge of histone modifications provides more precise prediction for enhancers and has been widely used in many studies. Even though there is no universal rule about which histone modifications should be used to identify enhancers, H3K4me1 and H3K27ac are well accepted as they are highly correlated with enhancer activity. This is also confirmed by advanced computational methods which use machine-learning approaches such as Hidden-Markov-Model (HMM) and Random-Forest to combine the signals of dozens of histone modifications to predict cis-regulatory elements including enhancers (Ernst and Kellis, 2012; Hoffman et al., 2012; Whalen et al., 2016). In addition, other datasets are also often integrated into the prediction, such as DNA accessibility, co-factor binding, RNA polymerase II binding and enhancer RNAs.

One mechanism by which enhancers can control distal genes is through their long-range physical interactions with promoters (Ong and Corces, 2011). It was proposed that the presence of enzymes at active promoters could also induce an opening of chromatin at enhancer regions because of the physical contacts between enhancers and promoters. This is supported by the finding of enhancer RNA (eRNA), a class of short non-coding RNA molecules that are transcribed from the DNA sequence of enhancer regions. Some potential functional roles of eRNA have been reported for regulating specific genes (Ilott et al., 2014; Kaikkonen et al., 2013, 2013, 2013; Li et al., 2013; Mousavi et al., 2013; Pnueli et al., 2015; Sigova et al., 2015).

Enhancer-promoter interaction networks are highly dependent on the nature of the genes. For instance, housekeeping genes seem to have no or few enhancers, whereas tissue-specific genes are often regulated by one or a few enhancers. Also, several enhancers can regulate one single promoter to fine tune the expression levels of their target gene in different cell types. Alternatively, one single enhancer can control several genes, which are usually coregulated (Andersson et al., 2014; Heinz et al., 2015; Li et al., 2012; Zhang et al., 2013).

Recently, enhancers with a special status called latent enhancers have been determined during stimulation in differentiated cells (Ostuni et al., 2013). These enhancers are not labeled by any enhancer-associated histone modifications (that is, H3K4me1 and H3K27ac) but acquire these active marks and transcription factor binding upon stimulation of cellular signaling pathways. This suggests an epigenetic memory of the exposure to environmental agents in cis-regulatory repertoire.

Another interesting form of enhancers is called super-enhancer, which is a region of the mammalian genome comprising a cluster of enhancers that is collectively bound by high level of transcription factors (Pott and Lieb, 2015; Whalen et al., 2016; Whyte et al., 2013). Super-enhancers are frequently identified near genes important for controlling and defining cell identity, and can therefore be used to quickly identify key nodes determining cell identity.

To validate the prediction of enhancers, the DNA sequence of an enhancer can be directly tested for its ability to enhance the transcription from a minimal core promoter. This includes traditional approaches such as the luciferase assay and genome-wide approaches like Self-transcribing active regulatory region sequencing (STARR-seq) (Arnold et al., 2013).

Polycomb-group proteins

Polycomb-group proteins are a family of proteins first discovered in *Drosophila* as genes important for embryonic development (Di Croce and Helin, 2013). Polycomb-group proteins are well known for silencing *Hox* genes through modulation of chromatin configuration.

Many Polycomb-group proteins are conserved in eukaryotes and form at least two distinct complexes, the Polycomb-repressive complexes 1 and 2, PRC1 and PRC2, respectively.

PRC1 and PRC2 silence genes either synergistically or independently (van Kruijsbergen et al., 2015). In the classical model, PRC2 is involved in recruiting PRC1 to the promoters of their common target genes (Boyer et al., 2006; Cao et al., 2002; Wang et al., 2004). In mammals, EZH2 is an important PRC2 subunit, which is a histone methyltransferase (HMT) that deposits H3K27me3 via its SET domain. The histone modification H3K27me3 is a hallmark of gene repression and usually found in the promoter regions of developmental genes. This modification can be recognized by a family of proteins which contain the chromobox (CBX) domain and function as subunits of PRC1, thereby recruit PRC1 to the H3K27me3 regions set by PRC2. Subsequently, PRC1 catalyzes the monoubiquitylation of histone H2A at lysine 119 (H2AK119ub) by the two ring-domain-containing proteins, RING1B and BMI1. Recent studies have demonstrated that PRC1 can also be involved in the recruitment of PRC2 by its deposition of H2AK119ub (Blackledge et al., 2014; Cooper et al., 2014). Of note, it is also shown that PRC1 can still deposit H2AK119ub and repress gene transcription independently of PRC2 and H3K27me3 (Tavares et al., 2012). In addition, DNA methylation is also an important determinant of PRC2 recruitment (Reddington et al., 2013).

Polycomb group proteins have long been linked to cell-fate specification (Burton and Torres-Padilla, 2014; Sparmann and van Lohuizen, 2006). In embryonic stem cells, lineage-specific genes, which define the identities of cells during differentiation, are primed for expression and kept in a repressed state known as bivalent state by the histone modification H3K4me3 and H3K27me3 together with PRC1 and PRC2. Dysregulation of Polycomb protein levels often results in early embryonic lethality (Sauvageau and Sauvageau, 2010). Polycomb group proteins have also been linked to the progression of cancer, as malfunction of these proteins can lead to activation of proto-oncogenes and the loss of function of tumor suppressor genes (Bracken and Helin, 2009; Sauvageau and Sauvageau, 2010; Sparmann and van Lohuizen, 2006).

3D genome — high-level structure of chromosome

Individual regulatory DNA elements and genes are not instructive for operating and directing a cell without the communications between them and their interactions with other molecules. Hence, it is not surprising that gene regulation relies not only on the local chromatin configuration, but also on the three-dimensional (3D) conformation of chromosomes which compartmentalizes the nucleus and facilitates the communication between separated functional DNA elements especially the CREs and their target genes. The genome together with a large variety of small molecules and biological macromolecules are housed in the

heart of a eukaryotic cell, the nucleus. However, the diameter of a nucleus is hundreds of thousands folds smaller than the total length of the genome. This discrepancy implicates a complex hierarchical compaction and organization of the chromosomes that is deployed to fit the genome into the nucleus while still ensuring the accessibility of desired sequences. Such hierarchical organization can be classified into at least three layers: primary structure, comprising DNA sequences, DNA-binding proteins, and nucleosomes that determine which regulatory DNA elements are accessible to transcription factors (TFs); secondary structure, comprising interactions between nearby nucleosomes that define local chromatin architecture; and tertiary structure, comprising long-range chromatin interactions that can span even more than a few hundreds of kilo bases and facilitate the communications between distal regulatory DNA elements and their target genes. All the above layers make the nucleus a crowded environment yet well-organized network that administrates the communication between diverse cellular elements.

Active and repressive chromatin

The genome can be divided into two main compartments which represent the two states of its secondary structure, being active and repressive. These two compartments are displayed in a cell-type specific manner and can be defined using different approaches such as open chromatin assays, ChIP-seq for a combination of histone modifications or profiling of lamina-associated domains (LADs) and nucleolus-associated domains (NADs) (Pombo and Dillon, 2015).

A recent technique called Hi-C provides an additional way to classify the compartments based on the long-range physical chromatin interaction patterns in the genome. The two main compartments identified by Hi-C study are termed A and B compartments which are associated with active and repressive chromatin respectively. Genomic loci from the same compartment type preferentially interact with each other rather than to the loci from the opposite compartment (Dixon et al., 2015; Lieberman-Aiden et al., 2009). Moreover, A/B compartments coincide genome-wide with the active and repressive chromatin identified by other independent assays for the same cells such as profiling of histone marks, DNA replication timing, LADs and NADs (Dixon et al., 2015; Rao et al., 2014). One further study using deep sequenced Hi-C data revealed that at least six subcompartments can be observed correlating with distinct patterns of histone modifications (Rao et al., 2014). Such patterns indicate relative nuclear positioning of individual regions of subcompartments.

Topologically associating domains

The building blocks of a genome are not randomly arranged in the nucleus but hierarchically organized rendering a similar structure as the chapters and paragraphs of a book. Firstly, in eukaryotic nuclei, individual chromosomes are mostly aggregated in their own territories.

This is supported by both microscopy-based fluorescence in situ hybridization (FISH) (Bolzer et al., 2005; Meaburn and Misteli, 2007; Parada et al., 2004) and sequencing-based whole genome derivatives of 3C experiments (Kalhor et al., 2012; Lieberman-Aiden et al., 2009). Secondly, the chromosomes are further partitioned into a number of relatively discrete regions termed topologically associating domains (TADs) (Dixon et al., 2012; Rao et al., 2014). A TAD is a continuous genomic region that constrains the physical interactions to occur mainly inside the fragment of relatively isolated chromatin. Depending on the resolution, the size of TADs range from a few hundreds of kilo bases to several millions of bases. Regions delimiting TADs are called TAD boundaries which usually contain insulator sites and are enriched with housekeeping genes (Dixon et al., 2012; Rao et al., 2014; Van Bortle et al., 2014).

TADs are conserved among different cell types and behave as the basic units of the genome. During DNA replication, replication domain boundaries coincide nearly one-to-one with TAD boundaries, suggesting TADs are stable regulatory units of replication timing (Dileep et al., 2015; Kalhor et al., 2012). It has also been shown that the regions displaying switching of A/B compartments during stem cell differentiation typically correspond to single or series of TADs, indicating that TADs are the units of dynamic alterations in chromosome compartments (Dixon et al., 2015). Furthermore, the identities of TADs in replication-timing and nuclear compartmentalization are significantly associated with distinctive patterns of epigenetic marks (Rao et al., 2014).

At TAD boundaries, the CCCTC-binding factor (CTCF) has been identified as one of the key players which together with cohesin are involved in the formation of a loop between two distal genomic loci (Phillips and Corces, 2009). However, many genomic CTCF-binding sites are not located at TAD boundaries but within TADs even with a 3D map of the human genome at kilobase resolution, indicating that CTCF and cohesin alone are insufficient to separate different TADs (Dixon et al., 2012; Rao et al., 2014). This is also supported by the mild effects on TAD organization in the cells that are depleted of CTCF or cohesion (Seitan et al., 2013; Sofueva et al., 2013; Zuin et al., 2014). How the TADs are created remains unclear.

Chromatin interactions between DNA regulatory elements

The cells with large genomes have evolved mechanisms to compartmentalize chromosomes at all scales, which provide more precise control of interactions mainly between certain regulatory DNA elements. Genome-wide maps of the proteins that bind to enhancers, promoters, and insulators, together with high resolution maps of the physical chromatin contacts, revealed that each chromosome contains thousands of DNA loops which are associated with CTCF and its looping partner cohesion (Dixon et al., 2012, 2015; Rao et al., 2014; Tang et al., 2015). These architectural loops further compartmentalize TADs into smaller insulated domains which ensure the correct wiring of enhancers and promoters.

Several lines of evidence suggest that enhancers regulate any gene that is co-localized within the same insulated domain (Andersson et al., 2014; Symmons et al., 2014). Although most chromatin interactions between regulatory elements are restricted within the same domain, many inter-chromosomal interactions were also detected and validated to be functionally important (Joshi et al., 2015; Schoenfelder et al., 2015; Spilianakis and Flavell, 2004; Zhang et al., 2013).

OUTLINE OF THE THESIS

Effects of chemical signals on chromatin organization

One fascinating aspect of cells is that they can receive and process signals (mostly chemical signals) from outside in order to respond to the changes in their living environment. In multicellular organisms, cells use many classes of chemical signals, including hormones and cytokines, to regulate physiology and behavior of their surrounding cells or target distal organs or tissues. Such regulations are compiled by the binding of signaling molecules to specific receptor proteins in the target cells resulting in a change in cell function. It has been shown that such functional modulations of cells happens as early as at the transcriptional level (Barolo and Posakony, 2002). However, mechanisms of chemical-signal-mediated transcriptional response are still poorly understood, especially at the level of chromatin topology.

Accurate regulation of transcription is mediated by interaction networks consisting of promoters, enhancers and many proteins including transcription factors. During chemical stimulation, receptor proteins recognize the particular structure of the signaling molecules, and then trigger the corresponding cellular biological pathway in which transcription factors (TFs) are involved to alter the expression of target genes. It is interesting that genome-wide studies have revealed many TFs bind not only to promoter regions, but often to genomic regions that are distal to promoters (Carroll et al., 2006; Consortium, 2012; John et al., 2011; Nielsen et al., 2008; Rao et al., 2011; Welboren et al., 2009). Such observation implicates that spatial organization of chromatin is likely associated with chemical-signal-mediated transcriptional response.

To investigate the effects of chemical signals on chromatin reorganization, we set out to study the effect of external stimuli by triamcinolone acetonide (TA) and tumor necrosis factor alpha (TNFa) on gene expression and chromatin organization. TA, which belongs to a class of steroid hormones called glucocorticoids (GCs), can activate glucocorticoid receptor (GR) pathway and regulate genes controlling development, metabolism, and immune response (Kadmiel and Cidlowski, 2013). TNFa, a proinflammatory cytokine, can activate nuclear factor kappa-b

(NF-κB) pathway which regulates genes influencing a broad range of biological processes including innate and adaptive immunity, inflammation and stress response (Hoesel and Schmid, 2013). Both GR and NF-κB can rapidly act in response to stimuli as they reside in the cytoplasm and remain inactive in their corresponding protein complexes, and do not require synthesis of new protein in order to become activated. GR is dissociated from the complex upon the binding of glucocorticoid and subsequently translocates into the nucleus where it binds to chromatin to regulate target genes (Kadmiel and Cidlowski, 2013). Similar to GR, NF-κB is released upon stimuli, translocates to the nucleus, and subsequently regulates target genes via its binding to chromatin (Hoesel and Schmid, 2013). Activated GR and NFκB predominantly bind to genomic regions that are distal to promoters. Only around 25% of the binding sites contain DNA motifs for GR and NFκB respectively, indicating other proteins facilitate the association of GR and p65 with chromatin. Indeed, P300 is present at almost all of these binding sites, which is a co-factor shared by GR and NFκB and generally considered as a hallmark of active enhancers (Rao et al., 2011). Co-activated GR and NFκB share a large proportion of binding sites and correlate with expression of many genes in a mutual antagonistic or synergistic manner. However, it was not entirely clear how GR and NFκB regulate their target genes from their distal binding sites.

Dynamics of chromatin organization between two states of pluripotency

Embryonic stem cells (ESCs) are able to differentiate and generate all cell types in the body. This ability is termed pluripotency. The pluripotent and self-renewal properties of ESCs make them an invaluable model for fundamental research into the regulatory mechanisms in early development. ESCs are derived from the inner cell mass (ICM) of a blastocyst and classically cultured in growth media supplemented with fetal calf serum and leukemia Inhibitory Factor (LIF). More recently, ESCs are also cultured in serum-free 2i medium that contains LIF plus two small-molecule kinase inhibitors PD0325901 and CHIR99021, targeting the mitogen-activated protein kinase (MEK) pathway and the glycogen synthase kinase-3 (GSK3) pathway, respectively (Ying et al., 2008). It is well accepted that 2i ESCs represent a ground-state pluripotency, whereas the classical serum-derived ESCs are reminiscent of post-implantation pluripotent stem cells (Marks and Stunnenberg, 2014; Marks et al., 2012; Nichols and Smith, 2009; Plusa and Hadjantonakis, 2014; Ying et al., 2008).

Both mouse embryonic stem cells (mESCs) grown in serum+LIF (‘‘serum mESCs’’) and 2i+LIF media (‘‘2i mESCs’’) are pluripotent; however, they show distinct transcriptomic profiles and epigenetic landscapes (Habibi et al., 2013; Marks and Stunnenberg, 2014; Marks et al., 2012). 2i mESCs have higher expression levels of metabolic genes and diminished expression levels of lineage-priming genes compared to serum mESCs. In serum mESCs, the promoters of around 3,000 genes display a bivalent chromatin state (co-occurrence of Histone marks H3K27me3

and H3K4me3) and are said to be poised for activation; these genes predominantly are involved in cell-fate determination and development. In 2i mESCs, the number of bivalent promoters drastically reduces; however, loss of repressive mark H3K27me3 does not result in transcriptional activation. In addition, the genome of 2i mESCs is globally hypo-methylated similar to the cells in ICM, whereas serum mESCs are hyper-methylated. Additional data from our lab show that only minor changes in H3K4me3 and H3K9me3 were observed between serum and 2i mESCs, but enhancer related marks H3K4me1 and H3K4me3 are undergoing extensive redistribution. However, it has not been explored whether 3D genomic organization play a role in initiation and/or maintenance of the two states of pluripotency by mediating the physical communication between regulatory DNA elements.

Epigenomic reprogramming during monocyte-to-macrophage differentiation and 'Innate immune memory'

The immune system in vertebrates protects organisms from infection and is categorized into two layers, innate immunity and adaptive immunity.

Innate immunity refers to nonspecific defense mechanisms that come into play immediately or within hours of infection or injury in the body, and is found in all classes of plants and animals. Upon infection, innate immune system initiates a process called inflammation to promote healing of damaged tissue following the clearance of pathogens. Various cell types involved in such processes including resident macrophages, dendritic cells, or blood circulating monocytes and neutrophils. These cells detect infection or injury using pattern recognition receptors (PRRs) on the surface or within the cell. PRRs recognize two classes of molecules. One class refers to pathogen-associated molecular patterns (PAMPs) which are broadly shared by pathogens but distinguishable from host molecules, including viral nucleic acids, fungal beta-glucan, and lipopolysaccharide (LPS) from Gram-negative bacteria. The other class refers to damage-associated molecular patterns (DAMPs) that are released from injured cells. Dysregulation of inflammation might lead the body to a state like sepsis or autoimmunity.

Immunological memory is a phenomenon that after an initial response to a specific pathogen, the immune system responds with the same pathogen more strongly to subsequent encounters, which is the basis of vaccination. The consensus was that immunological memory is an exclusive characteristic of adaptive immunity in vertebrates as innate immune system mounts only nonspecific responses. The discovery of PRRs challenged this dogma. Moreover, recent studies have shown a similar phenomenon in the absence of adaptive immune system in plants, invertebrates and also vertebrates, yet the second exposures were different from the first one. These studies suggest a trained immunity of innate immune system which can

be primed by a first stimulus and display enhanced responses upon later infection that can be maintained for up to a few months. Currently, the molecular bases of trained immunity are partially defined, but evidence supports the convergence of multiple regulatory layers, including changes in chromatin organization and the persistence of microRNAs (miRNAs) induced by the primary stimulus.

Monocytes and macrophages play an important role in the pathophysiology of sepsis and inflammation, along with other innate and adaptive immune cells (Biswas and Lopez-Collazo, 2009). Stimulating monocytes or macrophages with LPS inhibits their functions in immune response and causes so called LPS-induced tolerance in which these cells fail to properly response to re-stimulation. Transcriptome analysis of tolerant monocytes from sepsis patients (Shalova et al., 2015) and a mouse sepsis model (Foster et al., 2007) reveal that the tolerized phenotype cannot be explained purely through failure of specific signaling pathways induced by PRRs to activate downstream genes. This implicates a role for local chromatin architecture and specific transcriptional regulators in controlling the expression of tolerized genes (Glass and Natoli, 2016). Studies in human cancers have revealed commonalities between inflammation and cancer associated tolerance, including the role for IDO1 in both (Bessede et al., 2014). Accordingly, several anti-cancer drugs, such as bromodomain and extraterminal domain family (BET) inhibitors (Nicodeme et al., 2010) and a topoisomerase inhibitor (Rialdi et al., 2016) have proven efficacious in blocking inflammation-associated death in mice. The specific epigenetic and transcriptional remodeling induced by the initial LPS exposure, and the extent to which it specifies tolerance to future LPS exposure, is an important question.

IN THIS THESIS

In this chapter, **Chapter 1**, I briefly introduced the basic concepts related to our studies, including essential knowledge of epigenetic mechanisms and current understandings of 3D genomic organization.

In **Chapter 2**, we investigated the impact of signal-dependent transcription factors on gene regulation and chromatin reorganization. We determined epigenetic dynamics and TF binding patterns in the cells of which glucocorticoid receptor and nuclear factor kappa-b pathways are activated by glucocorticoid (GC) and tumor necrosis factor alpha (TNFa) respectively. We then investigated how those TF binding sites influenced gene expression via chromatin interactions using circular chromosome conformation capture coupled with next generation sequencing (4C-seq) and high-resolution chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) of P300. This study suggests an active role of signal-dependent transcription factors in chromatin and long-range interaction remodeling. In this study, I analyzed the data of ChIA-PET, 4C-seq, and ChIP-seq, and helped drafting the manuscript.

In **Chapter 3**, we investigated the dynamics of 3D genomic organization during the transition between two states of pluripotency. We established Capture Hi-C with target-sequence enrichment of DNase I hypersensitive sites. We observed a group of extremely long-range intra- and inter-chromosomal interactions (ELRIs) between a small subset of H3K27me3 marked bivalent promoters involving the *Hox* clusters in serum-grown but not in 2i-grown mESCs. We then investigated the role of PRC2 in establishing these interactions using Capture Hi-C on *Eed*$^{-/-}$ serum mESCs. Our study together with a parallel study suggest that PRC2 likely acts as an initiator of ELRIs by deposition of H3K27me3 and subsequent recruitment of canonical PRC1, which may act as the physical mediator of ELRIs. In this study, I analyzed the data of Capture HiC, 4C-seq, DNase I-seq, ChIP-seq, and RNA-seq, and helped drafting the manuscript.

In **Chapter 4**, we integrated time-resolved epigenomes and transcriptomes to characterize the molecular events involved in LPS-induced tolerance. We investigated epigenetic and transcriptional changes during LPS stimulated, beta-glucan stimulated and normal monocyte-to-macrophage differentiation. We further investigated epigenetic and transcriptional changes in tolerized macrophages under secondary LPS exposure. In addition, we assessed the ability of beta-glucan exposure to tolerance reversal both *in vitro* and *in vivo*. Our findings show that the innate immune ‘‘training stimulus’’ beta-glucan can reverse macrophage tolerance ex vivo and move forward in understanding how the tolerized phenotype can be reversed in sepsis patients. In this study, I helped in the analysis of RNA-seq and ChIP-seq data and drafting the manuscript.

In **Chapter 5**, we performed Capture Hi-C to investigate the dynamics of 3D genomic organization during monocyte-to-macrophage differentiation and immunological training. We investigated the associations between the changes at epigenetic, transcriptional and 3D genomic organization levels. We demonstrate that despite the absence of cell division, the monocyte-to-macrophage differentiation is associated with higher-order chromatin reorganization, including changes in A/B compartments, strengthening of TADs and their boundaries and a global gain of short-range interactions. While differentiation is associated with profound changes, immunological training does not cause spatial chromatin dynamics, suggesting that differentiation is the main driver of chromatin remodeling. In this study, I analyzed the data of Capture Hi-C, ChIP-seq, and RNA-seq, and helped drafting the manuscript.

REFERENCES

1. Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature 507, 455–461.
2. Arnold, C.D., Gerlach, D., Stelzer, C., Boryń, Ł.M., Rath, M., and Stark, A. (2013). Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq. Science *339*, 1074–1077.
3. Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K., and Schapira, M. (2012). Epigenetic protein families: a new frontier for drug discovery. Nat. Rev. Drug Discov. *11*, 384–400.
4. Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.
5. Bannister, A.J., and Kouzarides, T. (2005). Reversing histone methylation. Nature *436*, 1103–1106.
6. Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.
7. Barolo, S., and Posakony, J.W. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. Genes Dev. *16*, 1167–1181.
8. Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation. Cell *143*, 470–484.
9. Bartolomei, M.S., and Ferguson-Smith, A.C. (2011). Mammalian Genomic Imprinting. Cold Spring Harb. Perspect. Biol. *3*, a002592.
10. Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. Nat. Rev. Genet. *12*, 554–564.
11. Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y., Noushmehr, H., Lange, C.P.E., van Dijk, C.M., Tollenaar, R.A.E.M., et al. (2012). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat. Genet. *44*, 40–46.
12. Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas III, E.J., Gingeras, T.R., et al. (2005). Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse. Cell *120*, 169–181.
13. Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The Mammalian Epigenome. Cell *128*, 669–681.
14. Bessede, A., Gargaro, M., Pallotta, M.T., Matino, D., Servillo, G., Brunacci, C., Bicciato, S., Mazza, E.M.C., Macchiarulo, A., Vacca, C., et al. (2014). Aryl hydrocarbon receptor control of a disease tolerance defence pathway. Nature *511*, 184–190.
15. Biswas, S.K., and Lopez-Collazo, E. (2009). Endotoxin tolerance: new mechanisms, molecules and clinical significance. Trends Immunol. *30*, 475–487.
16. Black, J.C., Van Rechem, C., and Whetstine, J.R. (2012). Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. Mol. Cell *48*, 491–507.

17. Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. Cell *157*, 1445–1459.
18. Blackwood, E.M., and Kadonaga, J.T. (1998). Going the Distance: A Current View of Enhancer Action. Science *281*, 60–63.
19. Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., et al. (2005). Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes. PLOS Biol. *3*, e157.
20. Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349–353.
21. Bracken, A.P., and Helin, K. (2009). Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat. Rev. Cancer *9*, 773–784.
22. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.
23. Burton, A., and Torres-Padilla, M.-E. (2014). Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. Nat. Rev. Mol. Cell Biol. *15*, 723–735.
24. Calo, E., and Wysocka, J. (2013). Modification of Enhancer Chromatin: What, How, and Why? Mol. Cell *49*, 825–837.
25. Calvanese, V., Lara, E., Kahn, A., and Fraga, M.F. (2009). The role of epigenetics in aging and age-related diseases. Ageing Res. Rev. *8*, 268–276.
26. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. Science *298*, 1039–1043.
27. Carroll, J.S., Meyer, C.A., Song, J., Li, W., Geistlinger, T.R., Eeckhoute, J., Brodsky, A.S., Keeton, E.K., Fertuck, K.C., Hall, G.F., et al. (2006). Genome-wide analysis of estrogen receptor binding sites. Nat. Genet. *38*, 1289–1297.
28. Chen, Z., and Riggs, A.D. (2011). DNA Methylation and Demethylation in Mammals. J. Biol. Chem. *286*, 18347–18353.
29. Choy, M.-K., Movassagh, M., Goh, H.-G., Bennett, M.R., Down, T.A., and Foo, R.S. (2010). Genome-wide conserved consensus transcription factor binding motifs are hyper-methylated. BMC Genomics *11*, 519.
30. Consortium, T.E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.
31. Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N.P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., et al. (2014). Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. Cell Rep. *7*, 1456–1470.

32. Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat. Rev. Genet. *14*, 390–403.
33. Denker, A., and Laat, W. de (2016). The second decade of 3C technologies: detailed insights into nuclear organization. Genes Dev. *30*, 1357–1382.
34. Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nat. Struct. Mol. Biol. *20*, 1147–1155.
35. Dileep, V., Ay, F., Sima, J., Vera, D.L., Noble, W.S., and Gilbert, D.M. (2015). Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program. Genome Res.
36. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.
37. Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.
38. Egger, G., Liang, G., Aparicio, A., and Jones, P.A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. Nature *429*, 457–463.
39. Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.
40. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.
41. Feil, R., and Fraga, M.F. (2012). Epigenetics and the environment: emerging patterns and implications. Nat. Rev. Genet. *13*, 97–109.
42. Foster, S.L., Hargreaves, D.C., and Medzhitov, R. (2007). Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature *447*, 972–978.
43. Fukuda, H., Sano, N., Muto, S., and Horikoshi, M. (2006). Simple histone acetylation plays a complex role in the regulation of gene expression. Brief. Funct. Genomic. Proteomic. *5*, 190–208.
44. Glass, C.K., and Natoli, G. (2016). Molecular control of activation and priming in macrophages. Nat. Immunol. *17*, 26–33.
45. Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: A Landscape Takes Shape. Cell *128*, 635–638.
46. Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. Nature *389*, 349–352.
47. Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells. Cell Stem Cell *13*, 360–369.

48. Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: The reference human genome annotation for The ENCODE Project. Genome Res. *22*, 1760–1774.
49. He, H.H., Meyer, C.A., Hu, S.S., Chen, M.-W., Zang, C., Liu, Y., Rao, P.K., Fei, T., Xu, H., Long, H., et al. (2014). Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. Nat. Methods *11*, 73–78.
50. Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat. Rev. Mol. Cell Biol. *16*, 144–154.
51. Heyn, H., and Esteller, M. (2012). DNA methylation profiling in the clinic: applications and challenges. Nat. Rev. Genet. *13*, 679–692.
52. Hoesel, B., and Schmid, J.A. (2013). The complexity of NF-κB signaling in inflammation and cancer. Mol. Cancer *12*, 86.
53. Hoffman, M.M., Buske, O.J., Wang, J., Weng, Z., Bilmes, J.A., and Noble, W.S. (2012). Unsupervised pattern discovery in human chromatin structure through genomic segmentation. Nat. Methods *9*, 473–476.
54. Ilott, N.E., Heward, J.A., Roux, B., Tsitsiou, E., Fenwick, P.S., Lenzi, L., Goodhead, I., Hertz-Fowler, C., Heger, A., Hall, N., et al. (2014). Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human monocytes. Nat. Commun. *5*, 3979.
55. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. Science *333*, 1300–1303.
56. John, S., Sabo, P.J., Thurman, R.E., Sung, M.-H., Biddie, S.C., Johnson, T.A., Hager, G.L., and Stamatoyannopoulos, J.A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat. Genet. *43*, 264–268.
57. Jones, P.A., Issa, J.-P.J., and Baylin, S. (2016). Targeting the cancer epigenome for therapy. Nat. Rev. Genet. *17*, 630–641.
58. Joshi, O., Wang, S.-Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell Stem Cell *17*, 748–757.
59. Kadmiel, M., and Cidlowski, J.A. (2013). Glucocorticoid receptor signaling in health and disease. Trends Pharmacol. Sci. *34*, 518–530.
60. Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., et al. (2013). Remodeling of the Enhancer Landscape during Macrophage Activation Is Coupled to Enhancer Transcription. Mol. Cell *51*, 310–325.
61. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat. Biotechnol. *30*, 90–98.
62. Kelly, T.K., De Carvalho, D.D., and Jones, P.A. (2010). Epigenetic modifications as therapeutic targets. Nat. Biotechnol. *28*, 1069–1078.

63. Khorasanizadeh, S. (2004). The Nucleosome: From Genomic Organization to Genomic Regulation. Cell *116*, 259–272.
64. Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.
65. van Kruijsbergen, I., Hontelez, S., and Veenstra, G.J.C. (2015). Recruiting polycomb to chromatin. Int. J. Biochem. Cell Biol. *67*, 177–187.
66. Li, G., Levitus, M., Bustamante, C., and Widom, J. (2005). Rapid spontaneous accessibility of nucleosomal DNA. Nat. Struct. Mol. Biol. *12*, 46–53.
67. Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. Cell *148*, 84–98.
68. Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.
69. Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.
70. Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat. Rev. Genet. *11*, 285–296.
71. Marks, H., and Stunnenberg, H.G. (2014). Transcription regulation and chromatin structure in the pluripotent ground state. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1839*, 129–137.
72. Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis Stewart, A., Smith, A., et al. (2012). The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. Cell *149*, 590–604.
73. Marmorstein, R., and Zhou, M.-M. (2014). Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. Cold Spring Harb. Perspect. Biol. *6*, a018762.
74. Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.
75. Meaburn, K.J., and Misteli, T. (2007). Cell biology: Chromosome territories. Nature *445*, 379–381.
76. Meyer, C.A., and Liu, X.S. (2014). Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. Nat. Rev. Genet. *15*, 709–721.
77. Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371–378.
78. Mousavi, K., Zare, H., Dell'Orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul, A., Hager, G.L., and Sartorelli, V. (2013). eRNAs Promote Transcription by Establishing Chromatin Accessibility at Defined Genomic Loci. Mol. Cell *51*, 606–617.
79. Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012a). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature *489*, 83–90.

80. Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012b). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature *489*, 83–90.
81. Nichols, J., and Smith, A. (2009). Naive and Primed Pluripotent States. Cell Stem Cell *4*, 487–492.
82. Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., et al. (2010). Suppression of inflammation by a synthetic histone mimic. Nature *468*, 1119–1123.
83. Nielsen, R., Pedersen, T.Å., Hagenbeek, D., Moulos, P., Siersbæk, R., Megens, E., Denissov, S., Børgesen, M., Francoijs, K.-J., Mandrup, S., et al. (2008). Genome-wide profiling of PPARγ:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. Genes Dev. *22*, 2953–2967.
84. Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293.
85. Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent Enhancers Activated by Stimulation in Differentiated Cells. Cell *152*, 157–171.
86. Parada, L.A., McQueen, P.G., and Misteli, T. (2004). Tissue-specific spatial organization of genomes. Genome Biol. *5*, R44.
87. Phillips, J.E., and Corces, V.G. (2009). CTCF: Master Weaver of the Genome. Cell *137*, 1194–1211.
88. Plank, J.L., and Dean, A. (2014). Enhancer Function: Mechanistic and Genome-Wide Insights Come Together. Mol. Cell *55*, 5–14.
89. Plongthongkum, N., Diep, D.H., and Zhang, K. (2014). Advances in the profiling of DNA modifications: cytosine methylation and beyond. Nat. Rev. Genet. *15*, 647–661.
90. Plusa, B., and Hadjantonakis, A.-K. (2014). Embryonic stem cell identity grounded in the embryo. Nat. Cell Biol. *16*, 502–504.
91. Pnueli, L., Rudnizky, S., Yosefzon, Y., and Melamed, P. (2015). RNA transcribed from a distal enhancer is required for activating the chromatin at the promoter of the gonadotropin α-subunit gene. Proc. Natl. Acad. Sci. *112*, 4369–4374.
92. Pombo, A., and Dillon, N. (2015). Three-dimensional genome architecture: players and mechanisms. Nat. Rev. Mol. Cell Biol. *16*, 245–257.
93. Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat. Biotechnol. *28*, 1057–1068.
94. Pott, S., and Lieb, J.D. (2015). What are super-enhancers? Nat. Genet. *47*, 8–12.
95. Rao, N.A.S., McCalman, M.T., Moulos, P., Francoijs, K.-J., Chatziioannou, A., Kolisis, F.N., Alexis, M.N., Mitsiou, D.J., and Stunnenberg, H.G. (2011). Coactivation of GR and NFKB alters the repertoire of their binding sites and target genes. Genome Res. *21*, 1404–1416.
96. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

97. Reddington, J.P., Perricone, S.M., Nestor, C.E., Reichmann, J., Youngson, N.A., Suzuki, M., Reinhardt, D., Dunican, D.S., Prendergast, J.G., Mjoseng, H., et al. (2013). Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. Genome Biol. *14*, R25.
98. Rialdi, A., Campisi, L., Zhao, N., Lagda, A.C., Pietzsch, C., Ho, J.S.Y., Martinez-Gil, L., Fenouil, R., Chen, X., Edwards, M., et al. (2016). Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation. Science aad7993.
99. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.
100. Saeed, S., Quintin, J., Kerstens, H.H.D., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.-C., Ratter, J., Berentsen, K., Ent, M.A. van der, et al. (2014). Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science *345*, 1251086.
101. Sauvageau, M., and Sauvageau, G. (2010). Polycomb Group Proteins: Multi-Faceted Regulators of Somatic Stem Cells and Cancer. Cell Stem Cell *7*, 299–313.
102. Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., et al. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat. Genet. *47*, 1179–1186.
103. Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res.
104. Seto, E., and Yoshida, M. (2014). Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. Cold Spring Harb. Perspect. Biol. *6*, a018713.
105. Shalova, I.N., Lim, J.Y., Chittezhath, M., Zinkernagel, A.S., Beasley, F., Hernández-Jiménez, E., Toledano, V., Cubillos-Zapata, C., Rapisarda, A., Chen, J., et al. (2015). Human Monocytes Undergo Functional Re-programming during Sepsis Mediated by Hypoxia-Inducible Factor-1α. Immunity *42*, 484–498.
106. Sharp, A.J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S.B., Dupre, Y., and Antonarakis, S.E. (2011). DNA methylation profiles of human active and inactive X chromosomes. Genome Res. *21*, 1592–1600.
107. Sigova, A.A., Abraham, B.J., Ji, X., Molinie, B., Hannett, N.M., Guo, Y.E., Jangi, M., Giallourakis, C.C., Sharp, P.A., and Young, R.A. (2015). Transcription factor trapping by RNA in gene regulatory elements. Science *350*, 978–981.
108. Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.
109. Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. Nature *511*, 611–615.

110. Sofueva, S., Yaffe, E., Chan, W.-C., Georgopoulou, D., Rudan, M.V., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J. *32*, 3119–3129.
111. Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat. Rev. Cancer *6*, 846–856.
112. Spilianakis, C.G., and Flavell, R.A. (2004). Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat. Immunol. *5*, 1017–1027.
113. Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. *20*, 267–273.
114. Stunnenberg, H.G., Abrignani, S., Adams, D., de Almeida, M., Altucci, L., Amin, V., Amit, I., Antonarakis, S.E., Aparicio, S., Arima, T., et al. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell *167*, 1145–1149.
115. Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.
116. Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. Genome Res. *24*, 390–400.
117. Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science *324*, 930–935.
118. Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell *163*, 1611–1627.
119. Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. Cell *148*, 664–678.
120. Tessarz, P., and Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. Nat. Rev. Mol. Cell Biol. *15*, 703–708.
121. Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.
122. Van Bortle, K., Nichols, M.H., Li, L., Ong, C.-T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. Genome Biol. *15*, R82.
123. Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., et al. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. Genome Res. *23*, 555–567.
124. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The Sequence of the Human Genome. Science *291*, 1304–1351.

125. Wang, L., Brown, J.L., Cao, R., Zhang, Y., Kassis, J.A., and Jones, R.S. (2004). Hierarchical Recruitment of Polycomb Group Silencing Complexes. Mol. Cell *14*, 637–646.

126. Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. Genes Dev. *28*, 672–682.

127. Welboren, W.-J., Driel, M.A. van, Janssen-Megens, E.M., Heeringen, S.J. van, Sweep, F.C., Span, P.N., and Stunnenberg, H.G. (2009). ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. EMBO J. *28*, 1418–1428.

128. Whalen, S., Truty, R.M., and Pollard, K.S. (2016). Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat. Genet. *48*, 488–496.

129. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

130. Wu, H., and Zhang, Y. (2014). Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. Cell *156*, 45–68.

131. Yao, L., Berman, B.P., and Farnham, P.J. (2015). Demystifying the secret mission of enhancers: linking distal regulatory elements to target genes. Crit. Rev. Biochem. Mol. Biol. *50*, 550–573.

132. Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

133. Yu, M., Hon, G.C., Szulwach, K.E., Song, C.-X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., et al. (2012). Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. Cell *149*, 1368–1380.

134. Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev. *25*, 2227–2241.

135. Zhang, Y., Wong, C.-H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E., et al. (2013). Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature *504*, 306–310.

136. Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.

137. Zhu, H., Wang, G., and Qian, J. (2016). Transcription factors as readers and effectors of DNA methylation. Nat. Rev. Genet. *17*, 551–565.

138. Zuin, J., Dixon, J.R., Reijden, M.I.J.A. van der, Ye, Z., Kolovos, P., Brouwer, R.W.W., Corput, M.P.C. van de, Werken, H.J.G. van de, Knoch, T.A., IJcken, W.F.J. van, et al. (2014). Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc. Natl. Acad. Sci. *111*, 996–1001.

Table 1

Assay abbreviation	Full assay name	Refs
1 versus 1		
3C	Chromosome conformation capture	(Dekker et al., 2002)
1 versus Many/All		
4C	Circular chromosome conformation capture	(Zhao et al., 2006)
4C	Chromosome conformation capture-on-chip	(Simonis et al., 2006)
e4C	Enhanced chromosome conformation capture-on-chip	(Schoenfelder et al., 2010)
*Multiplex*ed 3C-seq	Multiplexed chromosome conformation capture sequencing	(Ribeiro de Almeida et al., 2011)
4C-seq	Chromosome conformation capture-on-chip combined with high-throughput sequencing	(Splinter et al., 2011)
Many versus Many		
5C	Chromosome conformation capture carbon copy	(Dostie et al., 2006)
ChIA–PET	Chromatin interaction analysis paired-end tag sequencing	(Fullwood et al., 2009)
Many versus All		
Capture-3C	Chromosome conformation capture coupled with oligonucleotide capture technology	(Hughes et al., 2014)
Capture-HiC	Hi-C coupled with oligonucleotide capture technology	(Dryden et al., 2014)
All versus All		
Hi-C	Genome-wide chromosome conformation capture	(Lieberman-Aiden et al., 2009)
ELP	Genome-wide chromosome conformation capture with enrichment of ligation products	(Tanizawa et al., 2010)
TCC	Tethered conformation capture	(Kalhor et al., 2012)
Single-cell Hi-C	Single-cell genome-wide chromosome conformation capture	(Nagano et al., 2013)
In situ Hi-C	Genome-wide chromosome conformation capture with ***in situ*** ligation	(Rao et al., 2014)
DNase Hi-C	Genome-wide chromosome conformation capture with DNase I digestion	(Ma et al., 2015)
Micro-C	Genome-wide chromosome conformation capture with micrococcal nuclease digestion	(Hsieh et al., 2015)

References:

1. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing Chromosome Conformation. Science ***295***, 1306–1311.
2. Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. Genome Res. ***16***, 1299–1309.
3. Dryden, N.H., Broome, L.R., Dudbridge, F., Johnson, N., Orr, N., Schoenfelder, S., Nagano, T., Andrews, S., Wingett, S., Kozarewa, I., et al. (2014). Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. Genome Res. ***24***, 1854–1868.
4. Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-α-bound human chromatin interactome. Nature ***462***, 58–64.
5. Hsieh, T.-H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell ***162***, 108–119.
6. Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat. Genet. ***46***, 205–212.
7. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat. Biotechnol. ***30***, 90–98.
8. Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science ***326***, 289–293.
9. Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., et al. (2015). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. Nat. Methods ***12***, 71–78.
10. Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature ***502***, 59–64.
11. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell ***159***, 1665–1680.
12. Ribeiro de Almeida, C., Stadhouders, R., de Bruijn, M.J.W., Bergen, I.M., Thongjuea, S., Lenhard, B., van IJcken, W., Grosveld, F., Galjart, N., Soler, E., et al. (2011). The DNA-Binding Protein CTCF Limits Proximal Vκ Recombination and Restricts κ Enhancer Interactions to the Immunoglobulin κ Light Chain Locus. Immunity ***35***, 501–513.
13. Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., et al. (2010). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat. Genet. ***42***, 53–61.
14. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). Nat. Genet. ***38***, 1348–1354.
15. Splinter, E., Wit, E. de, Nora, E.P., Klous, P., Werken, H.J.G. van de, Zhu, Y., Kaaij, L.J.T., IJcken, W. van, Gribnau, J., Heard, E., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. Genes Dev. ***25***, 1371–1383.
16. Tanizawa, H., Iwasaki, O., Tanaka, A., Capizzi, J.R., Wickramasinghe, P., Lee, M., Fu, Z., and Noma, K. (2010). Mapping of long-range associations throughout the fission yeast genome reveals global genome organization linked to transcriptional regulation. Nucleic Acids Res. ***38***, 8164–8177.
17. Zhao, Z., Tavoosidana, G., Sjölinder, M., Göndör, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Singh Sandhu, K., Singh, U., et al. (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat. Genet. ***38***, 1341–1347.

Chapter 2

Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization

Kuznetsova *et al. Genome Biology* (2015) 16:264
DOI 10.1186/s13059-015-0832-9

Genome Biology

RESEARCH Open Access

Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization

Tatyana Kuznetsova[1†], Shuang-Yin Wang[1†], Nagesha A. Rao[1†], Amit Mandoli[1], Joost H. A. Martens[1], Nils Rother[1], Aafke Aartse[1], Laszlo Groh[1], Eva M. Janssen-Megens[1], Guoliang Li[2], Yijun Ruan[3], Colin Logie[1] and Hendrik G. Stunnenberg[1*]

Abstract

Background: The impact of signal-dependent transcription factors, such as glucocorticoid receptor and nuclear factor kappa-b, on the three-dimensional organization of chromatin remains a topic of discussion. The possible scenarios range from remodeling of higher order chromatin architecture by activated transcription factors to recruitment of activated transcription factors to pre-established long-range interactions.

Results: Using circular chromosome conformation capture coupled with next generation sequencing and high-resolution chromatin interaction analysis by paired-end tag sequencing of P300, we observed agonist-induced changes in long-range chromatin interactions, and uncovered interconnected enhancer–enhancer hubs spanning up to one megabase. The vast majority of activated glucocorticoid receptor and nuclear factor kappa-b appeared to join pre-existing P300 enhancer hubs without affecting the chromatin conformation. In contrast, binding of the activated transcription factors to loci with their consensus response elements led to the increased formation of an active epigenetic state of enhancers and a significant increase in long-range interactions within pre-existing enhancer networks. De novo enhancers or ligand-responsive enhancer hubs preferentially interacted with ligand-induced genes.

Conclusions: We demonstrate that, at a subset of genomic loci, ligand-mediated induction leads to active enhancer formation and an increase in long-range interactions, facilitating efficient regulation of target genes. Therefore, our data suggest an active role of signal-dependent transcription factors in chromatin and long-range interaction remodeling.

Keywords: ChIA-PET, Chromosome conformation capture, Enhancer, GR, Long-range interaction, NFκB, P300, Transcription regulation

Background

Mechanisms of transcriptional response mediated by signal-dependent transcription factors (inducible TFs) are not well understood at the level of chromatin topology. Recent genome-wide studies have revealed that the majority of TF binding sites (up to 90 %) are distal to promoters and located in intragenic and intergenic regions [1–9]. These studies collectively revealed cell-type-specific constellations of distal regulatory regions that change during differentiation and development in a highly ordered fashion, whereby some distal regulatory regions are being set up de novo and others are decommissioned. This implies that at least some lineage-specific and/or signal-dependent TFs effectively open the chromatin structure and prepare the chromatin for subsequent binding of other TFs. A simplistic model of how such plasticity can be achieved is that long-range interactions among and between enhancers and promoters are dynamically established or disrupted. Many recent studies have purported an active or "instructive" role of inducible TFs in mediating long-range chromatin contacts for efficient regulation of target genes [10–14]. The orchestrated long-range interaction changes have also been reported in embryonic stem cells (ESC) and

* Correspondence: h.stunnenberg@ncmls.ru.nl
†Equal contributors
[1]Department of Molecular Biology, Faculty of Science Nijmegen, Radboud University, Nijmegen, The Netherlands
Full list of author information is available at the end of the article

ESC-derived lineages [15] on a topological domain (TAD) level. In contrast, other studies suggest a static or "permissive" model in which inducible TFs passively join pre-existing interaction networks of regulatory elements without affecting the organization of long-range interactions [14, 16, 17]. The *HoxD* locus serves as an example of pre-formed long-range interactions [18]. Interestingly, in another report focusing on the *HoxD* locus, the authors directly compared the interaction profiles obtained by chromosome conformation capture (3C)-based methods and fluorescent in situ hybridization. The authors conclude that interactions identified by 3C-based methods at such high resolution do not always represent true proximal ligations, but may be a consequence of indirect cross-linking [19]. Discrepancies between studies on inducible TF-mediated long-range chromatin contacts may be due to differences in resolution and methodology or to the use of asynchronous cells.

Glucocorticoid receptor (GR) is a ligand inducible TF that belongs to the nuclear receptor superfamily [20]. Hormone binding dissociates the GR-containing cytoplasmic complex; GR then translocates to the nucleus where it binds to chromatin to regulate target gene activity. Nuclear factor kappa-b (NFκB) is a heterodimeric TF that regulates various biological processes such as cell growth, development, and the inflammatory response. In response to inflammatory stimuli such as the pro-inflammatory cytokine tumor necrosis factor alpha (TNFα), NFκB dissociates from an inhibitory cytoplasmic complex, translocates to the nucleus, and subsequently regulates its target genes [21–25]. Co-activated GR and NFκB share a large proportion of genomic regulatory elements and co-regulate many genes in a mutual antagonistic or synergistic manner [7, 26–29]. The majority of GR and p65 (a major NFκB subunit) binding events occur at genomic loci that exhibit pre-existing enhancer signatures. In this scenario, TFs other than GR and NFκB have established and maintain an open chromatin conformation, facilitating binding or recruitment of GR and p65 to their binding sites [30–32]. At a minority of GR and p65 binding sites (~10 %), the activated TFs establish de novo enhancer-like loci [5, 33, 34].

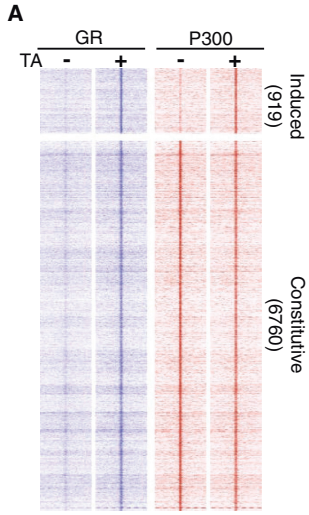
To gain insight in how GR and NFκB regulate their target gene repertoire from distal binding sites (DBSs), we mapped the chromatin interactions before and after GR and NFκB activation by generating high-resolution chromatin interaction profiles using the chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) method [35, 36]. We used antibodies against enhancer-associated P300 and against RNA polymerase II (POLII). P300 is a co-factor shared by GR and NFκB and its genomic occupancy in general is considered a hallmark of active enhancers [37–40]. We scrutinized the local chromatin interaction networks at genomic loci that are de novo established and compared them to those of pre-existing loci. We extended our analysis using high-resolution circular chromosome conformation capture (4C) technology on a subset of genomic viewpoints harboring de novo programmed regulatory elements. Collectively, our comprehensive analyses reveal a role of signal-dependent TF-induced dynamic changes in chromatin regulatory networks and its impact on gene regulation.

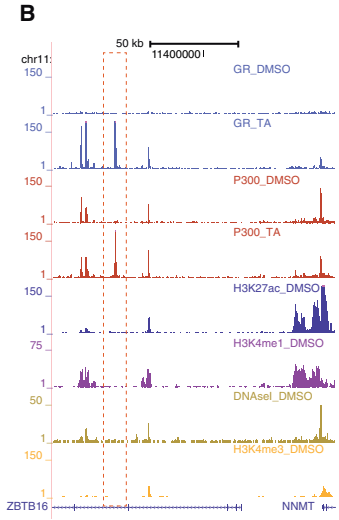
Results

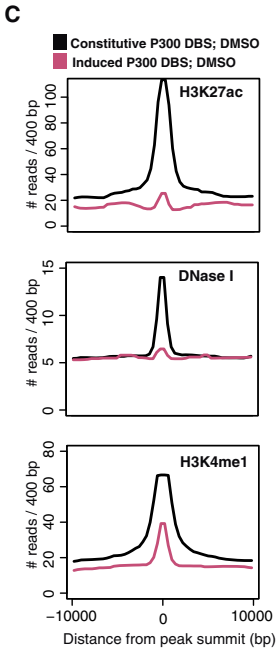
P300 is recruited to latent distal binding sites by ligand activated GR and/or NFκB

To gain insight into the impact of GR activation on the chromatin state and three-dimensional (3D) organization, we first performed chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) for GR, P300, epigenetic marks (H3K27ac, H3K4me3, and H3K4me1) and DNase I accessibility analysis. Ligand-activated GR binds to several thousand genomic loci [5, 7, 8, 41], of which more than 90 % (7679/8303) were located distally (>5 kb) from transcription start sites in HeLa cells (Additional file 1: Figure S1A). The vast majority (6760/7679) of these DBSs were DNase I accessible, bound by P300, and marked with H3K27ac and H3K4me1 prior to hormone stimulation (Fig. 1a–c). We refer to these as "pre-existing" P300 sites. Importantly, a subset of GR DBSs (919/7679) displayed the hallmarks of poised enhancers prior to ligand treatment, being largely inaccessible to DNase I, lowly marked with H3K4me1, and not marked with H3K27ac and P300. Interestingly, P300 was robustly recruited to these epigenetically dormant loci upon GR induction ("induced" P300 sites; Fig. 1a–c). At a smaller subset (529/6760) of GR DBSs, P300 occupancy was moderately reduced upon hormone treatment (data not shown). Next, we analyzed the pre-existing and induced P300 DBSs for TF motifs [7]. As expected, ligand-induced P300 DBS were highly enriched for glucocorticoid response elements (GREs), whereas AP1 was the most prevalent motif detected at pre-existing P300 sites (Fig. 1d). Our observations together with published data [5] suggest that at the induced P300 DBSs that are pre-marked with H3K4me1, GR binds directly to consensus GREs and recruits P300 to set up enhancer-like elements.

P65 was also predominantly bound at distal genomic loci (11,454/12,546) (Additional file 2: Figure S2A), of which the majority (10,453/11,454) were occupied by P300 prior to TNFα stimulation (pre-existing P300 sites). At a subset of p65 DBSs (1001/11,454), P300 was detectable only upon TNFα stimulation (Additional file 2: Figure S2B,C). TNFα-induced P300 DBSs were enriched for the NFκB response element (NFκB-RE) (Additional file 2: Figure S2D). Furthermore, induced P300 DBSs







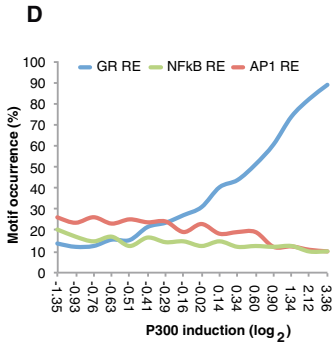


Fig. 1 (See legend on next page.)

(See figure on previous page.)
Fig. 1 Activated glucocorticoid receptor (*GR*) recruits P300 to epigenomically latent genomic regions. **a** Color profile depicting the GR and P300 signal at all GR-bound regions with either constitutive or ligand (triamcinolone acetonide [*TA*]) -induced P300 occupancy. **b** Example screenshot depicting the TA-induced P300 distal binding site (DBS; *dotted box*) and constitutive P300 DBS. **c** Basal (untreated cells) H3K27ac, DNase I hypersensitive site, and H3K4me1 signal at all GR-induced and constitutive P300 DBSs. **d** Motif occurrence at all GR-bound DBSs presented as a function of TA-dependent P300 recruitment (x-axis). *DMSO* dimethyl sulfoxide, *RE* response elements

that were barely or not marked by H3K27ac were inaccessible to DNase I, yet displayed readily detectable levels of H3K4me1 prior to TNFα induction and p65 binding (Additional file 2: Figure S2E). In line with a recent study in mouse macrophages [34], we presume that TNFα induction activates poised or latent enhancers. We also observed pre-existing P300 binding at many sites (~25,000) that were not significantly co-occupied by GR or p65 (Additional file 3: Figure S3A,B). These sites likely have a regulatory role in association with other TFs.

Because GR and p65 share a large number of regulatory elements (~30 %) and co-regulate many genes, we performed a similar analysis upon co-activation of GR and p65. We detected all the induced P300 DBSs that were uncovered upon single activation of GR or p65. An additional subset of inducible P300 sites (~700) was unveiled only upon co-stimulation, displaying significantly increased DNase I accessibility and H3K27ac, and a marginal increase in H3K4me1 (Additional file 4: Figure S4, Additional file 5: Figure S5, Additional file 6).

Taken together, GR and p65 mostly join pre-existing enhancer-like P300 DBSs that are set up by other TFs such as AP1. At a subset of latent genomic locations marked with low levels of H3K4me1, GR and/or NFκB binding induces DNA accessibility, recruitment of P300, and H3K27ac deposition. Because the induced P300 sites are highly enriched for their respective consensus response elements, it appears that recruitment of GR and NFκB to their respective cis-acting elements can initiate the formation of an active enhancer configuration, in line with recent studies [5, 33, 34].

ChIA-PET reveals P300 enhancer interaction networks

Next we focused on long-range chromatin contacts associated with P300 DBSs. We performed chromatin interaction analyses on co-stimulated cells to uncover the largest number of induced P300 DBSs (2881), and contrasted them to vehicle treatment. We performed ChIA-PET, an antibody-based method, to map the genome-wide chromatin interactions at high resolution [35, 36, 42]. We mapped the chromatin interactions using P300 and POLII antibodies. Sequencing of the P300 ChIA-PET libraries yielded 36.7 and 18.2 million uniquely mapped paired-end tags (PETs) for vehicle and co-stimulated samples, respectively. Among these, 1.4 and 1.2 million reads were self-ligation PETs (defined as ligation endpoints or anchors less than 5 kb apart) accounting for 15,148 and 16,366 putative P300 binding sites in vehicle and co-treated libraries, respectively (Additional file 7: Table S1). The vast majority (>90 %) of these self-ligation PETs co-localized with the P300 binding sites identified by ChIP-seq (Additional file 8: Figure S6A). ChIP-seq binding sites with low signal strength were not detected as binding sites in ChIA-PET data sets (Additional file 8: Figure S6B). Therefore, we used ChIP-seq binding sites (identified from ~20 million unique reads) as anchors to identify high confidence chromatin contacts. Ligation PETs that had their anchors between 5 and 1000 kb from each other and co-localized with high confidence P300 ChIP-seq binding sites were defined as long-range interactions. We identified 2363 and 5429 intra-chromosomal interactions using the P300 antibody in vehicle and co-stimulated cells, respectively. Using a similar approach, a large number of intra-chromosomal interactions were detected in a ChIA-PET analysis using a POLII antibody (Additional file 9: Table S2). P300 and POLII ChIP-seq binding sites that were involved in chromatin interactions were of higher signal strength compared to those not detected in chromatin interactions (Additional file 8: Figure S6C).

The majority of P300-associated long-range interactions occurred between distal regulatory elements (DBSs, ~60 %), whereas about 20 % occurred between promoters and DBSs (Fig. 2a). In contrast to the P300 interactome, POLII-associated interactions were found predominantly between promoters (64 %) and only 19 % involved DBS-promoter interactions (Fig. 2b). Visual inspection suggested that identified chromatin interactions occurred frequently between a multitude of P300 DBSs that aggregate into interaction subdomains (Additional file 8: Figure S6D) similar to replication or TADs [43, 44]. Indeed, more than 95 % of all P300 and POLII long-range interactions were confined to such domains as defined by DNA replication timing in HeLaS3 cells [44, 45] (Additional file 8: Figure S6E). Whereas the average TAD length is ~1.7 Mb, the average widths of P300 and POLII subdomains were 118 kb and 96 kb, respectively. Direct comparison of individual P300 and POLII interaction domains revealed that two fifths (39.6 %) overlapped, whereas the remainder appeared to involve only P300 or POLII (Fig. 2c, upper panel). The degree of P300 and POLII anchor overlap in

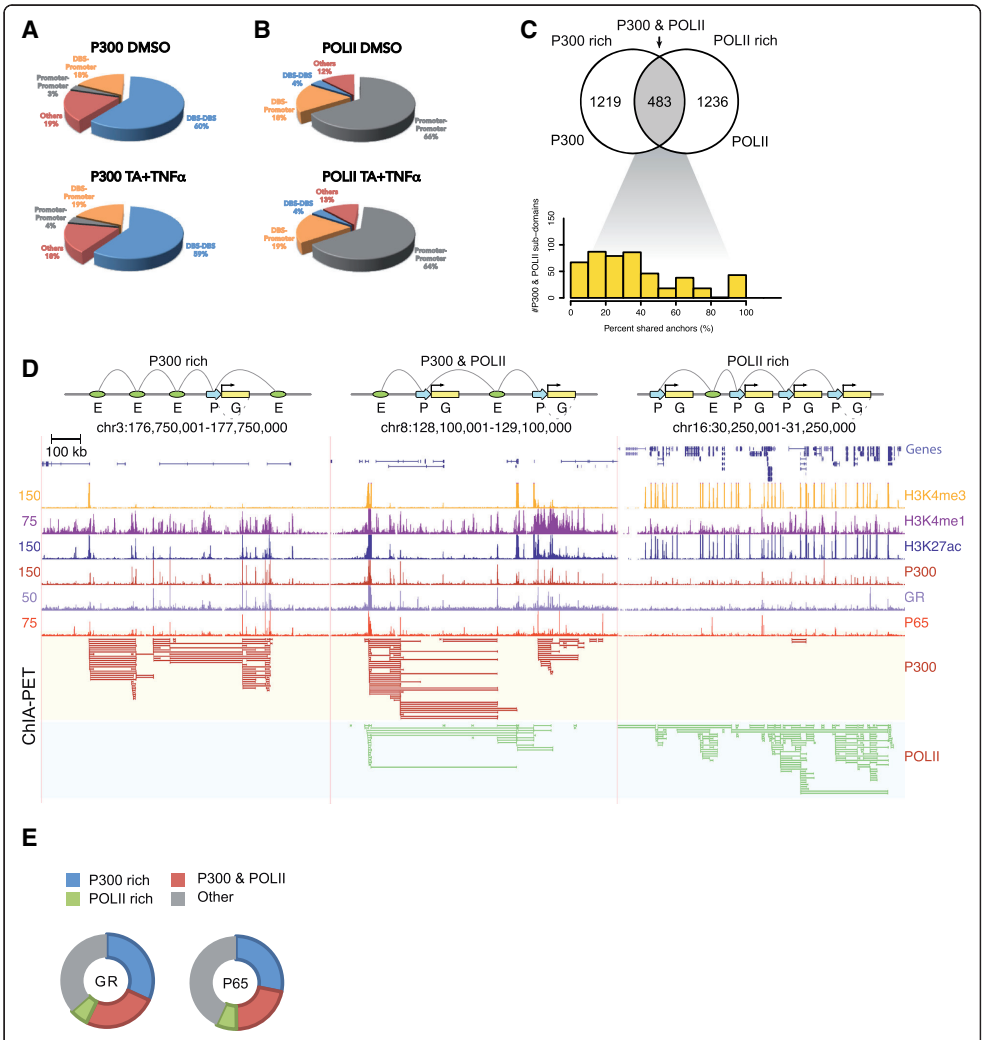


Fig. 2 P300 ChIA-PET interaction profile shows an enhancer-centered interaction pattern in contrast to promoter-centered RNA polymerase II (*POLII*) interactome. **a** Proportion of distal binding site (*DBS*)-promoter, promoter-promoter, and DBS-DBS interactions identified by P300 ChIA-PET in cells treated by vehicle (dimethyl sulfoxide [*DMSO*]) (*upper panel*) or triamcinolone acetonide (*TA*) + tumor necrosis factor alpha (*TNFα*) (*lower panel*). **b** Proportion of DBS-promoter, promoter–promoter, and DBS–DBS interactions identified by POLII ChIA-PET in cells treated with vehicle (*upper panel*) or TA + TNFα (*lower panel*). **c** Venn diagram depicting the extent of overlap between P300 interaction subdomains and POLII interaction subdomains (*upper panel*). Histogram depicting the percentage of P300 and POLII shared anchors in P300 & POLII interaction subdomains (*lower panel*). **d** Example screenshots of P300-rich (*left panel*), P300 & POLII-rich (*middle panel*), and POLII-rich (*right panel*) interaction subdomains depicting the ChIP-seq and ChIA-PET interaction data. **e** Distribution of glucocorticoid receptor (*GR*) (*left panel*) and p65 (*right panel*) binding sites in P300 rich, P300 & POLII-rich, and POLII-rich interaction subdomains

P300 and POLII shared interaction domains varied, with most of the subdomains sharing less than 50 % of anchors (Fig. 2c, lower panel). Representative examples of P300-rich, P300 and POLII, and POLII-rich interaction subdomains are shown in Fig. 2d.

Because GR and p65 preponderantly bind to putative enhancers that are marked with P300, it would be expected for GR and p65 binding sites to be enriched within the P300 ChIA-PET interaction network. Indeed, about 60 % of GR and 50 % of p65 binding sites were located within P300 centric interaction subdomains (P300 rich, P300 & POLII). POLII-rich promoter–promoter networks were largely devoid of GR and p65 binding events (Fig. 2d,e).

Ligand treatment enhances long-range interactions at induced P300 distal binding sites

Next we set out to investigate whether pre-existing and induced P300 sites participate equally in long-range chromatin interactions. Upon ligand activation we observed a significant gain of DNaseI accessibility and active chromatin marks at induced P300 DBSs. We reason that these sites might have an increased interaction upon ligand activation.

To validate the P300-mediated long-range interactions and to gain insight into their frequency, we selected 4C viewpoints in eight different P300 interaction subdomains that encompassed 58 different genomic loci (anchors) in our ChIA-PET analysis. 4C-seq libraries from at least two independent biological replicas per viewpoint were sequenced to obtain more than 2 million high-quality, uniquely aligned reads (Fig. 3, Additional file 10: Figure S7, Additional file 11: Figure S8, Additional file 12: Figure S9). This sequencing depth is regarded adequate to map all ligation events within the viewpoint [46].

In the *ZBTB16/NNMT* locus (Fig. 3a), transcription of the *NNMT* gene was induced by co-stimulation. This locus contained one ligand-induced P300 binding site that also gained H3K27ac and DNase I accessibility (see also Fig. 1b) and three pre-existing P300 sites. Using one of the pre-existing P300 sites as the viewpoint in 4C experiments, we detected its interaction with other pre-existing P300 sites. Upon ligand activation, we observed the formation of novel interactions involving the ligand-induced P300 DBS as well as a general increase in the interaction signal at pre-existing sites. In ChIA-PET, we detected interactions between all the enhancers only upon ligand induction.

The *KLF6* locus encompassed multiple constitutive and four induced P300 binding sites (Fig. 3b). Upon stimulation, transcription of the *KLF6* gene was highly induced and multiple enhancers gained P300, H3K27ac, and DNase I accessibility. A GR-induced DBS was used as the viewpoint for 4C. In vehicle-treated cells, we detected weak 4C signals between the bait and surrounding pre-existing and induced P300 DBSs. These contacts were robustly increased upon co-stimulation. An additional six genomic viewpoints showed a similar increase in interaction frequencies and inclusion of induced P300 binding sites in the interaction network upon ligand induction (Additional file 11: Figure S8, Additional file 12: Figure S9).

To assess the interaction frequency at P300 DBSs, we divided the P300 DBSs that were detected in our 4C analysis (eight viewpoints) into induced and pre-existing. For each group we plotted the average of P300 ChIP-seq and 4C signal (reads per kilobase per million mapped reads [RPKM]) in control and stimulated cells. The constitutive P300 binding sites displayed a similar ChIP-seq and 4C signal pattern in vehicle-treated and ligand-treated cells (Fig. 3c). Importantly, induced P300 binding sites showed a significantly higher ($p < 0.001$, t-test) 4C signal in ligand-treated cells than the control cells (Fig. 3d).

GR and NFκB activation enhances long-range chromatin contacts

The 4C assays support the presence of long-range interaction networks among P300 DBSs. Furthermore, they uncovered a significant increase in contact frequency at induced but not at pre-existing P300 DBSs (Fig. 3c,d). To further investigate this difference, we divided the ChIA-PET interaction subdomains into two groups: subdomains containing only pre-existing P300 DBSs, and subdomains containing at least one induced P300 DBS. We then compared their interactome in the ChIA-PET profiles. However, to directly compare the two conditions, the immunoprecipitation-introduced bias inherent to ChIA-PET had to be taken into account. The ChIP step results in a restricted representation of the interactome. A possible confounding factor in ChIA-PET is that chromatin regions with a higher number of binding sites with high occupancy (RPKM) – that is higher local concentration of P300 – may be ChIPed with higher efficiency than regions with fewer binding sites and lower P300 occupancy.

In order to accurately compare the pre-existing and induced subdomains in untreated and co-stimulated ChIA-PET libraries, we first estimated the local P300 concentrations (average P300 signal) by summing up the RPKM values of P300 DBSs in ChIA-PET interaction subdomains harboring at least five P300 DBSs with a different degree of P300 induction (Fig. 4a). With few exceptions, the co-stimulation marginally affected the local concentration of P300 as compared to vehicle-treated cells (<2-fold) (Fig. 4b). Next, we selected subdomains that upon co-stimulation responded with no more

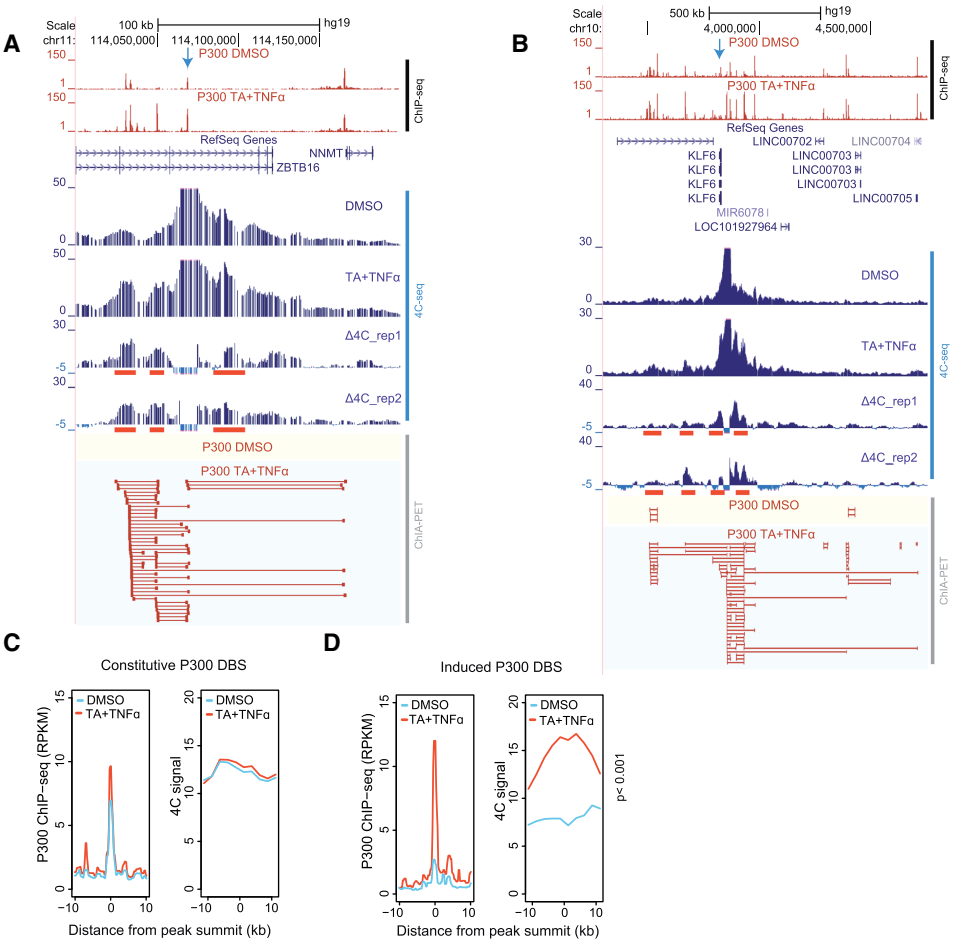


Fig. 3 Glucocorticoid receptor (GR)-dependent and p65 activation-dependent changes in chromatin interactions. **a**, **b** Direct comparison of chromatin contacts revealed by P300 ChIA-PET and 4C-seq analyses at the ZBTB16 (a) and KLF6 loci (b). 4C-seq bait loci are marked on each screenshot (*blue arrow*). Genomic regions that show a marked change in 4C signal upon triamcinolone acetonide (*TA*) + tumor necrosis factor alpha (*TNFα*) treatment are marked *red* below the Δ4C track. **c** Direct comparison of changes in average ChIP-seq signal (*left panel*) and 4C signal (*right panel*) at all the constitutive P300 distal binding sites (*DBSs*) within the ten 4C-seq genomic view points upon TA + TNFα treatment. **d** Direct comparison of changes in average ChIP-seq signal (*left panel*) and 4C signal (*right panel*) at all the induced P300 DBSs within the ten 4C-seq genomic view points upon TA + TNFα treatment. *DMSO* dimethyl sulfoxide

than a 25 % change in total P300 concentration (subdomains within the shaded area in Fig. 4b). We computed the chromatin interaction frequencies (ChIA-PET interactions/subdomain) in subdomains that had at least one or no induced P300 DBSs. In order to take into account the coverage difference of the two P300 ChIA-PET libraries, we used the one-sided Mann–Whitney–Wilcoxon test. Interestingly, P300 hubs bearing induced P300 DBSs displayed a significant increase in chromatin contacts upon co-stimulation versus vehicle-treated cells (Fig. 4c, upper panel). Such preference is not evident in subdomains harboring only pre-existing P300 DBSs (Fig. 4c, lower panel).

Thus, preferential involvement of ligand-induced P300 DBSs in chromatin interactions implies that GR and/or

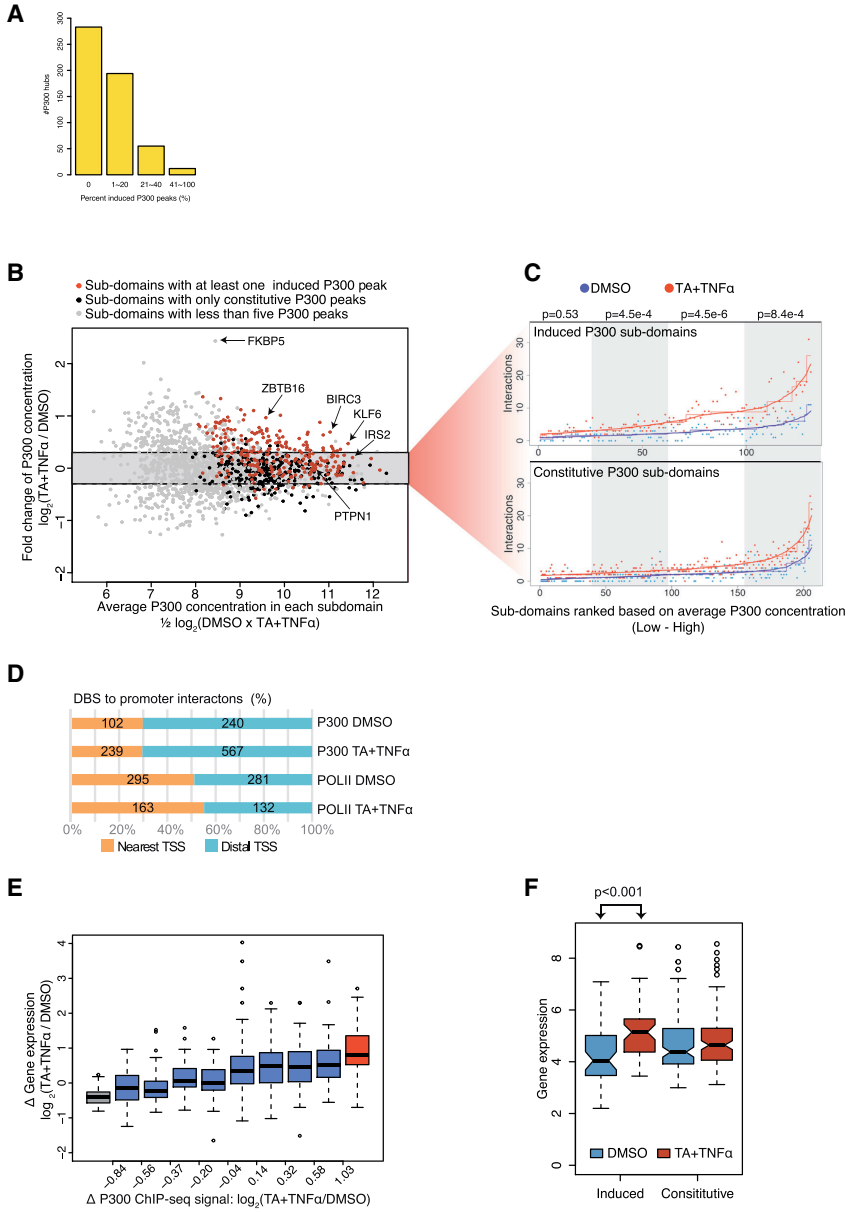


Fig. 4 (See legend on next page.)

(See figure on previous page.)
Fig. 4 Ligand-induced enhancement of chromatin contacts within P300 interaction subdomains. **a** Bar plot depicting the percentage of induced P300 peaks in the P300 hubs with at least five P300 peaks. **b** P300 local concentration at all P300 interaction subdomains that are ordered based on average P300 density (x-axis) and ligand-induced change in P300 concentration (y-axis). Sub-domains harboring at least five P300 distal binding sites (*DBS*) are presented**. c** Total number of chromatin interactions detected at each interaction subdomain upon vehicle dimethyl sulfoxide (*DMSO, blue dots*) and triamcinolone acetonide (*TA*) + tumor necrosis factor alpha (*TNFa, red dots*) treatment is presented. Sub-domains that show no more than 25 % change in total P300 concentration (reads per kilobase per million mapped reads) that either harbor at least one induced P300 DBS (*upper panel*) and only constitutive P300 DBS (*lower panel*) are used. *p*-values (Mann–Whitney) were generated by comparing interaction ratios (TA + TNFα/DMSO) between the induced and constitutive subdomains that are within each bin (shaded area). **d** Proportion of DBSs to nearest first gene promoters or DBSs to distal gene promoter interactions identified by P300 and RNA polymerase II (*POLII*) ChIA-PET analysis. *TSS* transcription start site. **e** Co-activation-dependent changes in expression of genes that are directly looped to P300 DBSs. P300 DBSs were initially ordered according to agonist-induced P300 recruitment (low to high) and subsequently divided into 10 equal bins. The average transcriptional change of genes in each bin is presented. Expression of genes that are in bins harboring significantly repressed (*gray*), induced (*red*), and constitutive (*blue*) P300 DBSs upon co-stimulation. **f** Co-activation-induced changes in expression of genes that are within interaction subdomains harboring either induced or only constitutive P300 DBSs

NFκB binding to DNA via their cognate cis-acting elements opens up the closed chromatin by recruiting chromatin-remodeling complexes. Such open chromatin regions preferentially contact other P300 DBSs with a similar regulatory factor composition and chromatin state, resulting in an interaction network that is synchronized upon ligand-dependent GR and/or NFκB recruitment, resulting in increased contact frequency. We preferentially detected the networks that were highly affected by ligand stimulation. These networks are already established prior to stimulation (as detected by 4C-seq). However, the increase of the contact frequency at induced P300 DBSs upon ligand activation indicates the synchronization of such networks.

We therefore consider that the (over) representation of induced interactions in our data sets may imply that a considerably larger proportion of cells in the population have the P300 protein network at these loci because the queried loci were synchronized by ligand treatment.

GR and NFκB networks are enriched with their target genes

Finally, we assessed the impact of agonist-induced regulatory elements and their special re-organization on transcriptional regulation using gene-body POLII density as a direct read-out [6]. We quantified all the DBS-promoter contacts that were identified in both P300 and POLII interaction data sets. Importantly, 70 % of P300 and 50 % of POLII bound DBS were not contacting the nearest (first) active transcription start site (TSS) but a more distal TSS (Fig. 4d). Next, we computed the gene-body POLII density of genes that were connected to (induced and pre-existing) P300 DBSs. We observed a consistent positive correlation between ligand-dependent gene induction and P300 induction at the DBS (Fig. 4e). Genes linked to induced P300 DBSs respond avidly to ligand induction as compared to genes that are interacting with pre-existing P300 DBSs. Similarly, expression of all genes in the agonist-induced interaction subdomains was significantly increased upon co-stimulation whereas genes in the constitutive subdomains were unresponsive to agonists (Fig. 4f). Gene ontology (GO) analysis of genes in the induced P300 hubs instigated by activation of NFκB (TNFα or co-stimulated) are enriched for GO terms associated with inflammatory response whereas genes linked to activated GR-induced P300 hubs are enriched for various biological processes, including macromolecule metabolic processes. Genes connected to constitutive P300 hubs are enriched for GO terms associated with general cellular processes (Additional file 13: Table S3). Taken together, our data strongly suggests that GRE-containing or NFκB RE-containing latent enhancers that are activated upon ligand stimulation preferentially engage the GR and NFκB target genes and subsequently modulate their expression.

Discussion

In this study, we have analyzed the impact of agonist-activated GR and p65 (NFκB) on chromatin state, 3D organization, and transcriptional regulation. Activated GR and p65 are mainly recruited to pre-existing regulatory elements that are pre-bound by P300, and display the epigenetic signature of active enhancers prior to TF activation, that is, they fall into the "permissive" model category. Our findings are in line with recent findings suggesting that signal-dependent TFs largely access the enhancer landscape that is set up by other lineage-specific TFs such as PU.1, C/EBPα, and AP1 [4, 5, 30, 34]. However, in ChIP-seq, we also observed de novo recruitment of P300 by activated GR and/or p65 to thousands of regions that subsequently acquired enhancer-like epigenetic features, in line with recent studies [5, 33, 34]. ChIA-PET and 4C analysis indicate that these sites are involved in interaction that would fall into the "instructive" category. The chromatin signature and epigenetic changes in response to ligand activation at these induced DBSs are reminiscent of "latent enhancers" [34, 47].

An important question is how the agonist-induced enhancers and their target genes are spatially organized. We have compared the P300-mediated chromatin interactomes before and after GR and NFκB co-activation. Using ChIA-PET against P300, we observed the formation of P300-mediated long-range interactions at subdomains bearing induced P300 DBSs in response to ligand activation. With the current depth of ChIA-PET libraries, these interactions appeared to be either formed de novo (from latent enhancers) or stabilized upon ligand induction. In 4C analysis we reproducibly detected increased interactions at induced P300 DBSs; the induction in the 4C approach is, however, less pronounced as compared to ChIA-PET. This difference is likely due to intrinsic differences between the two assays: 4C detects long-range interactions irrespective of the presence or absence of P300 and hence is able to detect lower strength or lower frequency interactions and those that are P300 independent. In ChIA-PET, the immunoprecipitation step enriches for interactions mediated by P300 and does not pick up P300 independent interactions, providing an all or nothing picture. Collectively and in agreement with each other, our ChIA-PET and 4C results show that activation of GR and/or NFκB facilitates an induced interaction signal at a subset of DBSs. We interpret this increased signal as an increase in either the interaction frequency (stabilization of a network) or in the proportion of cells that engage in such interaction (synchronization of a network).

Recent 3C-based studies of individual loci reported on the role of GR and NFκB in long-range gene regulation [10, 13, 48]. For example, the *Lcn2* gene locus is engaged in multiple long-range contacts with GR DBS. In agreement with our findings, it was shown that activated GR increases local chromatin interactions without dramatic change in 3D organization. In another report exploiting the 4C approach, activated GR was shown to bind a downstream enhancer of the *Tsc22d3* gene, causing a 2-fold increase in long-range enhancer–promoter interaction and activation of transcription [13]. Similarly, TNFα induces chromatin interactions between distal NFκB-bound enhancers and the promoter proximal regulatory sites of *CCL2* [10]. In contrast, a recent study based on genome wide Hi-C analysis revealed that the vast majority of TNFα responsive enhancers, as determined by p65 binding, show little change in DNA looping after TNFα treatment [16]. The authors note that only ~15 % of p65 DBSs display an activated enhancer signature (increase in H3K27ac signal and enhancer RNA production) upon TNFα treatment. The apparent discrepancy with our study is likely due to the differences in resolution of applied techniques. In agreement with Jin et al., we found that the majority of long-range interactions are pre-established and not dynamic; however, by applying ChIA-PET and high-resolution 4C, we found a significant increase in long-range interactions at induced but not constitutive P300 DBSs. These changes in a subset of interactions are conceivably difficult to pick up using a relatively low-resolution Hi-C-only approach.

One of the questions debated in the field of chromatin topology is the extent to which long-range interactions are dynamic and correlate with gene expression, such as in response to extracellular stimuli or during differentiation. The instructive model suggests de novo formation of long-range interactions, where lineage-specific and/or signal-induced TFs establish a new interaction landscape and affect the expression of their target genes [10–14, 49]. Our data provides support for this model: at ligand-induced DBSs, ChIA-PET and 4C data show an increased interaction signal at loci that were largely closed with low or no active epigenetic marking (H3K27ac) but with low levels of H3K4me1, reminiscent of latent enhancers [34]. Our data also provide support for the permissive model, showing that the long-range interaction landscape is pre-formed in the absence of ligand induction. Ligand-activated TFs appear to join a pre-set network of enhancers and trigger transcription by lineage-specific and/or signal-induced TFs [16–18].

Conclusions

We conclude that the ligand-activated GR and p65 induce chromatin accessibility, P300 recruitment, and alterations of 3D chromatin structure at a subset of genomic loci. At subdomains with induced P300 binding, activated GR and p65 facilitate close spatial proximity of the induced P300 DBS with a pre-existing interaction network and enhancement of 3D chromatin contacts. Our data suggest that ligand induction causes synchronization or stabilization of active chromatin states and higher order structure in a large proportion of cells to facilitate efficient regulation of their target genes. We speculate that this spatial clustering of regulatory elements can cause an increase in the local concentration of regulatory proteins, which ultimately can enhance the transcriptional activity of associated genes. Further experiments are needed to validate and expand on these findings, to elucidate the role of inducible TFs in long-range regulation, and to firmly establish that increased physical looping interaction indeed leads to increased transcription.

Methods

Cell culture

HeLa B2 cells were maintained as described [7]. Cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10 % charcoal stripped fetal calf serum for 72–96 h before subsequent treatment and/or harvesting. Cells were treated with either DMSO or 1 μM of TA (T6501, Sigma-Aldrich, St. Louis, Missouri, United

States) for 4 h with or without an additional treatment with 10 ng/mL TNFα (T0157, Sigma-Aldrich) for the last hour.

ChIP-seq

ChIP was performed according to standard protocol [50] with minor modifications. Paraformaldehyde (1 %) cross-linking was carried out for 10 min followed by the chromatin preparation as described earlier [7]. Nuclei were re-suspended in ChIP-incubation buffer at a concentration of 20×10^6 cells/mL and sheared (seven cycles with each cycle containing 10 s power on and 10 s interval) using Bioruptor®Plus (B01020001, Diagenode, Liege, Belgium). Sonicated chromatin equivalent of 4×10^6 cells was incubated with relevant antibody overnight at 4 °C. Antibodies against P300 (sc-585x, Santa Cruz Biotechnology, Inc., Dallas, Texas, United States), POLII (MMS-126R-500, Covance, Inc., Princeton, New Jersey, United States), H3K27ac (C15410196, Diagenode), H3K4me1 (C15410194, Diagenode), and H3K4me3 (C15410003, Diagenode) were used. ChIP-seq sample preparation and sequencing was performed according to manufacturer's instructions (Illumina, San Diego, California, United States) and essentially as described [6, 9, 51] (http://www.blueprint-epigenome.eu).

ChIP-seq data analysis

The image files generated by HiSeq2000 (Illumina) were processed to extract sequence data and the 36/42 bp tags were unambiguously mapped to the human genome (NCBI, hg19) using the bwa aligner, allowing at most one nucleotide mismatch. Reads were further directionally extended to 200 bp, corresponding to the original length of the DNA fragments used for sequencing. For each base pair in the genome, the number of overlapping sequence reads was determined, averaged over a 10 bp window, and visualized in the University of California Santa Cruz genome browser (http://genome-euro.ucsc.edu). ChIP-seq data sets were normalized as described [6, 7] in order to eliminate the differences caused by sequencing depth/mapping efficiency.

Detection of putative P300 and POLII binding sites was performed using MACS (version 1.4.2) [52] with the p-value $<10^{-9}$. Peaks identified by using each antibody in DMSO, TA, TNFα, and TA + TNFα were combined in a common pool and sequence tags were counted under each peak location (for each data set separately). Then we calculated the intensity ($\log_2$ RPKM) of peaks in each treatment. Binding sites that showed a significant change (median ± 2 × median absolute deviation; $p < 0.05$) in signal for P300 or POLII in a treatment compared to that in the vehicle-treated sample were regarded as dynamic binding sites. Published GR, p65, and POLII ChIP-seq data that were generated in an identical experimental setup in HeLa B2 cells [GEO: GSE24518] were used in this study.

DNase I-seq

DNase I libraries were prepared from DMSO-treated and TA + TNFα-treated Hela B2 cells as described (http://www.uwencode.org/protocols). In brief, 5×10^6 nuclei were isolated using Buffer A (15 mM NaCl; 60 mM KCl; 1 mM EDTA, pH 8.0; 0.5 mM EGTA, pH 8.0; 15 mM Tris–HCl, pH 8.0; 0.5 mM spermidine) supplemented with 0.06 % IGEPAL CA-630 detergent. DNase I treatment (60 units) was performed for 3 min and the reaction stopped with stop buffer (50 mM Tris–HCl, pH 8; 100 mM NaCl; 0.10 % SDS; 100 mM EDTA, pH 8.0; 1 mM spermidine; 0.3 mM spermine). The sample was further fractionated on 9 % sucrose gradient for 24 h at 25,000 rpm at 16 °C. Fractions containing DNA fragments smaller than 1 kb were purified and processed for sequencing according to the Illumina library preparation protocol. Normalized (read number equalized) DNase I data sets were used for the downstream analysis and visualization.

ChIA-PET library preparation

ChIA-PET libraries were prepared using the standard protocol [35, 36]. Chromatin preparation and ChIP enrichment using P300 and POLII antibodies were performed as described above. Briefly, chromatin captured on magnetic beads was trimmed (blunt end), phosphorylated on 5′ ends, then underwent biotinylated half-linker ligation. Chromatin complexes were then divided into two equal halves and two independent half-linker ligation reactions were performed using half-linkers A and B containing specific barcodes (linker-A TAAG; linker-B ATGT). Subsequently, chromatin complexes were eluted from the beads and two linker ligation aliquots were combined together for proximity ligation under diluted conditions. Subsequently, reverse cross-linked and purified circular DNA was digested using MmeI enzyme (the restriction site is encoded on the linker). Next, biotinylated DNA fragments were immobilized on M-280 streptavidin Dynabeads (Invitrogen, Carlsbad, California, United States) followed by adaptor ligation. The efficiency of the library preparation was evaluated by polymerase chain reaction (PCR) and subsequent gel electrophoresis. Next, each library characterized by adaptor-ligated DNA fragments carrying 20 bp of genomic DNA flanking the 36-bp linker sequence on either side was sequenced on a HISeq200 (Illumina). A typical sequencing run yielding 200 million single-end reads of 100-bp length was generated for each library.

ChIA-PET data analysis

The first 72 bp of each sequenced read carrying the complete ChIA-PET ligation product (linker plus genomic DNA) was taken for the further analysis after trimming the ends of each read. Subsequently, single-end

sequenced reads were split at the linker ligation junction (linkerA/B-|-linkerA/B) and flipped to make the data compatible (similar to paired-end sequencing reads) for the ChIA-PET data analysis pipeline [42]. The average distance between binding sites (P300 and POLII), identified based on the ChIA-PET self-ligation PETs and the binding sites identified by ChIP-seq, were examined to ascertain the reproducibility of binding sites by these methods. The binding sites identified by both methods were highly comparable, but a larger number of total binding sites were identified by ChIP-seq owing to the higher sequencing depth. Therefore, we used ChIP-seq binding sites as anchors to identify the intra-chromosomal and inter-chromosomal interaction PETs. True long-range interaction signals were distinguished from the non-specific technical interaction noise by using the method described earlier [42]. Briefly, interaction PETs having a PET count equal to two or more for P300 libraries and three or more for POLII libraries at a false discovery rate <0.05 were considered as high confidence interaction clusters. We used a 5 kb and 1 Mb genomic span as the lower and upper cutoff limits, respectively, to define the high confidence interaction PET data. Each interaction PET contained a pair of interacting anchors. Direct overlap (book end or 1 bp) of anchors of each cluster with that of other clusters was performed to identify interaction complexes or interaction subdomains. Hence, the interaction clusters were further collapsed in to interaction complexes/subdomains based on the interconnectivity of the PET clusters.

Identification of dynamic interactions using ChIA-PET data sets

To minimize the bias induced by local P300 concentration on the chromatin interactions detected by ChIA-PET, we analyzed the changes of interaction frequencies per each P300 ChIA-PET defined subdomain as follows. For each subdomain, we counted the number of P300 peaks and calculated the P300 concentration (average $\log_2$ RPKM) in DMSO-treated and TA + TNFα-treated samples. All the subdomains are ranked by the average P300 concentration of DMSO and TA + TNFα treatments. We discarded the subdomains with less than five P300 peaks, and separated subdomains with at least one induced P300 peak (261 subdomains) and the ones with only constitutive P300 peaks (283 subdomains). We plotted the number of interaction clusters identified in DMSO and TA + TNFα data sets separately for individual subdomains. The subdomains were further filtered by the fold change of P300 concentration (>−0.3 and <0.3) and this resulted in 131 and 206 subdomains, respectively. The Mann–Whitney test was adopted to investigate the agonist-induced change in average chromatin interaction frequencies in comparable groups of subdomains that harbor only constitutive P300 DBSs against those having at least one agonist-induced P300 DBS.

4C-seq library preparations

4C assays were performed as described previously [46] with minor modifications. Briefly, 10^7 cells were cross-linked for 10 min with 2 % paraformaldehyde, quenched with glycine, and lysed in 50 mL lysis buffer (50 mM Tris, pH 7.5; 150 mM NaCl; 5 mM EDTA; 0.5 % NP-40; 1 % TX-100; 1X protease inhibitors) for 30 min. Nuclei were then digested by DpnII enzyme followed by inactivation of the restriction enzyme by incubating at 65 °C for 20 min. The digested chromatin was subsequently ligated (circularized) overnight at 16 °C with 50 U T4 ligase. Ligated chromatin was then reverse cross-linked by incubating with proteinase K at 65 °C and the RNA was removed by additional incubation at 37 °C with RNase A. The purified DNA was further digested with a second restriction enzyme of choice (BfaI, MseI, or NlaIII) followed by circularization of the DNA. The 4C product was subsequently amplified with bait-specific inverse primers (Additional file 14: Table S4). From each 4C library, about 3200 or 800 ng DNA was amplified in multiple parallel PCR reactions containing 200 ng of DNA each, which were subsequently pooled and purified. Amplified bait-containing DNA fragments were ligated to NextFlex DNA barcoded adaptors (Bioo Scientific, Austin, Texas, United States). Adaptor-ligated DNA was purified by Agencourt AMPure XP purification system (Beckman Coulter, Brea, California, United States), PCR amplified (eight cycles), and sequenced single-end on the Illumina HiSeq2000 to obtain 50-bp-long reads.

4C-seq data analysis

To improve the mappability of the sequence reads, we generated a reduced genome by extracting the sequences flanking the DpnII sites (30 bp on each strand from the DpnII sites to downstream) based on build version hg19 of the human genome. Then we estimated the mappability of the extracted sequences (each strand separately) and only uniquely mappable DpnII sites were considered for downstream analysis.

All the reads from each library were parsed based on the bait-specific primer sequence and mapped to the reduced genome using bwa (version 0.6.2) with the default parameters. The mapping data of the individual libraries are summarized (Additional file 15: Table S5). We initially mapped each replicate library separately and merged the replicate libraries based on their quality. The 4C signal was calculated using a sliding window of 10 kb (±5 kb of a given DpnII site) and normalized to the total number uniquely mapped reads. Δ4C is the difference of 4C signal in each genomic bin (10 kb) between the normalized DMSO and TA + TNFα data sets.

Gene ontology analysis

GO analysis was performed using the DAVID web tool [53, 54]. Gene sets were analyzed for enriched GO terms (biological processes) compared to the human genome database as background. Fisher's exact test was used to identify significantly enriched GO terms.

Data availability

All the ChIP-seq, ChIA-PET, and 4C raw data files have been submitted to GEO database [GEO: GSE61911]. Previously published GR, p65, and POLII ChIP-seq data can be accessed via [GEO: GSE24518].

Ethical approval

No approvals were required for the study, which complied with all relevant regulations.

Additional files

Additional file 1: Figure S1. GR binds predominantly to distal binding sites. Pile-up heat map depicting the H3K4me1 and H3K4me3 signal around (±12 kb) all GR-bound promoters and DBSs. (PDF 272 kb)

Additional file 2: Figure S2. Activated p65 induces de novo P300 depositions to latent genomic loci. **(A)** Pile-up heat map depicting the H3K4me1 and H3K4me3 signal around (±12 kb) all p65-bound promoters and DBSs. **(B)** Pile-up heat map depicting the p65 and P300 signal at all p65-bound enhancers upon vehicle, DMSO **(–),** and TNFα (**+**) treatment. **(C)** Example screenshot depicting TNFα-induced P300 recruitment at genomic regions (red box) and recruitment of p65 at genomic loci that are pre-marked by P300. **(D)** Motif occurrence at all p65-bound DBS presented as a function of TNFα-dependent P300 recruitment (x-axis) (top-panel). Level of shared binding of p65 and other TFs at all p65-bound DBS, presented as a function of TNFα-dependent P300 recruitment (bottom panel). **(E)** Level of H3K27ac, DNase I hypersensitivity, and H3K4me1 at all p65-bound DBSs (induced and constitutive P300 sites). (PDF 909 kb)

Additional file 3: Figure S3. Large numbers of P300-bound loci are not co-occupied by GR or p65. **(A)** Average ChIP-seq signal of GR and P300 around (±10 kb) all the P300 binding sites that do not show a significant GR occupancy. **(B)** Average ChIP-seq signal of p65 and P300 around (±10 kb) all the P300 binding sites that do not show a significant p65 occupancy. (PDF 168 kb)

Additional file 4: Figure S4. Complex epigenetic changes induced by co-activated GR and p65. An example screenshot depicting the GR-dependent (blue box), p65-dependent (red box), and co-stimulation-dependent (black box) induced P300 DBSs and constitutive DBSs. Dynamic changes in DNase I accessibility, epigenetic modifications, and RNA-POLII activity on genes upon co-activation are noticeable characteristics of this locus. (PDF 236 kb)

Additional file 5: Figure S5. Co-activation of GR and p65 induces additional de novo P300 DBSs. **(A)** Pile-up heat map depicting signal of P300, GR, and p65 at three groups (p65 dependent, GR dependent, and co-stimulation dependent) of inducible P300 DBSs that were identified upon co-activation of GR and p65. Motif occurrence (%) at all three clusters of induced P300 DBSs (bar graphs). **(B)** H3K27ac, DNase I accessibility, and H3K4me1 signal at all induced P300 DBSs upon vehicle (DMSO) and co-treatments (TA + TNFα). **(C)** H3K27ac, DNase I accessibility, and H3K4me1 signal at all constitutive P300 DBSs upon vehicle (DMSO) and co-treatments (TA + TNFα). (PDF 1137 kb)

Additional file 6: Supplementary text and Supplementary references. Cross-talk between GR and NFκB leads to complex changes in chromatin landscape. (DOCX 73 kb)

Additional file 7: Table S1. Overview of sequenced and mapped reads of all the ChIA-PET libraries. (XLS 28 kb)

Additional file 8: Figure S6. (A) Histogram depicting the genomic proximity (localization) of P300 binding sites identified by ChIP-seq in relation to those identified by using ChIA-PET self-ligation PETs. Identical comparison is performed for both DMSO-treated and TA + TNFα-treated data sets**. (B)** P300 ChIP-seq signal at P300 binding sites commonly identified by ChIP-seq and ChIA-PET and those binding sites that were uniquely detected in the ChIP-seq data set. **(C)** P300 ChIP-seq signal at P300 binding sites that were either involved (anchor) or not involved (non-anchor) in long-range interaction as identified by ChIA-PET analysis. **(D)** An example screenshot depicting the P300 interaction subdomains, P300 ChIP-seq binding sites in relation to topological domains as defined by replication timing data (www.replicationdomain.org). **(E)** Localization of all the interaction subdomains identified by ChIA-PET analysis (P300 and POLII) in relation to topological domains. (PDF 1041 kb)

Additional file 9: Table S2. List of all significant long-range interactions identified by P300 and POLII ChIA-PET analysis. (XLS 2165 kb)

Additional file 10: Figure S7. Reproducibility of 4C-seq biological replicates at the *FKBP5* locus. (PDF 1094 kb)

Additional file 11: Figure S8. Direct comparison of long-range interactions identified by P300 ChIA-PET and 4C-seq analyses at *FKBP5*, *DUSP1*, and *BIRC3* loci. (PDF 1037 kb)

Additional file 12: Figure S9. Direct comparison of long-range interactions identified by P300 ChIA-PET and 4C-seq analyses at *PTPN1*, *SULF1* and *ZFHX3* loci. (PDF 1147 kb)

Additional file 13: Table S3. Overview of enriched GO terms for the genes that are in various types of interaction subdomains. (XLS 23 kb)

Additional file 14: Table S4. Overview of restriction enzyme combinations and inverse PCR primer pairs used for 4C-seq. (XLS 31 kb)

Additional file 15: Table S5. Overview of sequenced and mapped reads of 4C-seq libraries. (XLS 31 kb)

Abbreviations

bp: base pair; ChIP: chromatin immunoprecipitation; DBS: distal binding site; ESC: embryonic stem cells; GO: gene ontology; GR: glucocorticoid receptor; GRE: glucocorticoid response element; kb: kilobase; Mb: megabase; NFκB: nuclear factor kappa-b; PCR: polymerase chain reaction; PET: paired-end tag; POLII: RNA polymerase II; RPKM: reads per kilobase per million mapped reads; TA: triamcinolone acetonide; TAD: topologically associated domain; TF: transcription factor; TNFα: tumor necrosis factor alpha; TSS: transcription start site.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

TK designed and performed the 4C experiments with the help of AA, NR, and LG; SYW analyzed the data; NAR prepared the ChIA-PET, ChIP-seq, and DNase I libraries with the help of AM and JM; EMJM carried out the sequencing of the libraries; GL and YR assisted with ChIA-PET data analysis; TK, SYW, NAR, CL, and HGS designed the experiments and drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank Sadia Saeed and Lusy Handoko from the Department of Molecular Biology, Radboud University, Nijmegen, for sharing their knowledge on the DNase I accessibility and ChIA-PET protocols, respectively. Computational analysis was carried out on the Dutch national e-infrastructure with the support of SURF Foundation. This project was supported by the 'Nederlandse Organisatie voor Wetenschappelijk Onderzoek' (NWO; grant - 820.02.002 and 917.11.322).

Author details

[1]Department of Molecular Biology, Faculty of Science Nijmegen, Radboud University, Nijmegen, The Netherlands. [2]National Key Laboratory of Crop Genetic Improvement, College of Informatics, Huazhong Agricultural University, Wuhan, China. [3]The Jackson Laboratory for Genomic Medicine, and Department of Genetic and Development Biology, University of Connecticut, 400 Farmington Ave., Farmington, CT 06030, USA.

Received: 27 May 2015 Accepted: 11 November 2015
Published online: 01 December 2015

References

1. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489:57–74.
2. Barish GD, Yu RT, Karunasiri M, Ocampo CB, Dixon J, Benner C, et al. Bcl-6 and NF-kappaB cistromes mediate opposing regulation of the innate immune response. Genes Dev. 2010;24:2760–5.
3. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, et al. Genome-wide analysis of estrogen receptor binding sites. Nat Genet. 2006; 38:1289–97.
4. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38:576–89.
5. John S, Sabo PJ, Thurman RE, Sung M-H, Biddie SC, Johnson TA, et al. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat Genet. 2011;43:264–8.
6. Nielsen R, Pedersen TA, Hagenbeek D, Moulos P, Siersbaek R, Megens E, et al. Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. Genes Dev. 2008; 22:2953–67.
7. Rao NAS, McCalman MT, Moulos P, Francoijs K-J, Chatziioannou A, Kolisis FN, et al. Coactivation of GR and NFKB alters the repertoire of their binding sites and target genes. Genome Res. 2011;21:1404–16.
8. Reddy TE, Pauli F, Sprouse RO, Neff NF, Newberry KM, Garabedian MJ, et al. Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. Genome Res. 2009;19:2163–71.
9. Welboren W-J, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, et al. ChIP-seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J. 2009;28:1418–28.
10. Teferedegne B, Green MR, Guo Z, Boss JM. Mechanism of action of a distal NF-kappaB-dependent enhancer. Mol Cell Biol. 2006;26:5759–70.
11. Papantonis A, Kohro T, Baboo S, Larkin JD, Deng B, Short P, et al. TNFα signals through specialized factories where responsive coding and miRNA genes are transcribed. EMBO J. 2012;31:4404–14.
12. Hu Q, Kwon Y-S, Nunez E, Cardamone MD, Hutt KR, Ohgi KA, et al. Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. Proc Natl Acad Sci U S A. 2008;105:19199–204.
13. Stavreva DA, Coulon A, Baek S, Sung MH, John S, Stixova L, et al. Dynamics of chromatin accessibility and long-range interactions in response to glucocorticoid pulsing. Genome Res. 2015. gr.184168.114
14. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature. 2013;502:499–506.
15. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518:331–6.
16. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature. 2013;503:290–4.
17. Kocanova S, Kerr EA, Rafique S, Boyle S, Katz E, Caze-Subra S, et al. Activation of estrogen-responsive genes does not require their nuclear co-localization. PLoS Genet. 2010;6, e1000922.
18. Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, et al. A regulatory archipelago controls Hox genes transcription in digits. Cell. 2011;147:1132–45.
19. Williamson I, Berlivet S, Eskeland R, Boyle S, Illingworth RS, Paquette D, et al. Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. Genes Dev. 2014;28:2778–91.
20. Nuclear Receptors Nomenclature Committee. A unified nomenclature system for the nuclear receptor superfamily. Cell. 1999;97:161–3.
21. Baeuerle PA. Pro-inflammatory signaling: last pieces in the NF-κB puzzle? Curr Biol. 1998;8:R19–22.
22. Barnes PJ, Karin M. Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases. N Engl J Med. 1997;336:1066–71.
23. Hayden MS, Ghosh S. Shared principles in NF-kappaB signaling. Cell. 2008;132:344–62.
24. Karin M. New twists in gene regulation by glucocorticoid receptor: is DNA binding dispensable? Cell. 1998;93:487–90.
25. Vallabhapurapu S, Karin M. Regulation and function of NF-kappaB transcription factors in the immune system. Annu Rev Immunol. 2009;27:693–733.
26. Adcock IM, Nasuhara Y, Stevens DA, Barnes PJ. Ligand-induced differentiation of glucocorticoid receptor (GR) trans-repression and transactivation: preferential targeting of NF-kappaB and lack of I-kappaB involvement. Br J Pharmacol. 1999;127:1003–11.
27. Beck IME, Vanden Berghe W, Vermeulen L, Bougarne N, Vander Cruyssen B, Haegeman G, et al. Altered subcellular distribution of MSK1 induced by glucocorticoids contributes to NF-kappaB inhibition. EMBO J. 2008;27:1682–93.
28. Ito K, Barnes PJ, Adcock IM. Glucocorticoid receptor recruitment of histone deacetylase 2 inhibits interleukin-1beta-induced histone H4 acetylation on lysines 8 and 12. Mol Cell Biol. 2000;20:6891–903.
29. Ray A, Prefontaine KE. Physical association and functional antagonism between the p65 subunit of transcription factor NF-kappa B and the glucocorticoid receptor. Proc Natl Acad Sci U S A. 1994;91:752–6.
30. Biddie SC, John S, Sabo PJ, Thurman RE, Johnson TA, Schiltz RL, et al. Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. Mol Cell. 2011;43:145–55.
31. Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity. 2010;32:317–28.
32. Natoli G, Ghisletti S, Barozzi I. The genomic landscapes of inflammation. Genes Dev. 2011;25:101–6.
33. Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, et al. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol Cell. 2013;51:310–25.
34. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, et al. Latent enhancers activated by stimulation in differentiated cells. Cell. 2013;152:157–71.
35. Fullwood MJ, Han Y, Wei C-L, Ruan X, Ruan Y. Chromatin interaction analysis using paired-end tag sequencing. Curr Protoc Mol Biol. 2010;Chapter 21(Unit 21):15.1–25.
36. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature. 2009;462:58–64.
37. Chakravarti D, LaMorte VJ, Nelson MC, Nakajima T, Schulman IG, Juguilon H, et al. Role of CBP/P300 in nuclear receptor signalling. Nature. 1996;383:99–103.
38. Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, Collins T. CREB-binding protein/p300 are transcriptional coactivators of p65. Proc Natl Acad Sci U S A. 1997;94:2927–32.
39. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009;459:108–12.
40. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009;457:854–8.
41. De Bosscher K, Vanden Berghe W, Haegeman G. The interplay between the glucocorticoid receptor and nuclear factor-kappaB or activator protein-1: molecular mechanisms for gene repression. Endocr Rev. 2003;24:488–522.
42. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell. 2012;148:84–98.
43. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80.
44. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, et al. Topologically associating domains are stable units of replication-timing regulation. Nature. 2014;515:402–5.
45. Weddington N, Stuy A, Hiratani I, Ryba T, Yokochi T, Gilbert DM. ReplicationDomain: a visualization tool and comparative database for genome-wide replication timing data. BMC Bioinformatics. 2008;9:530.
46. van de Werken HJG, Landan G, Holwerda SJB, et al. Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat Meth. 2012;9:969–72.
47. Saeed S, Quintin J, Kerstens HHD, Rao NA, Aghajanirefah A, Matarese F, et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science. 2014;345:1251086.

48. Hakim O, Sung M-H, Voss TC, Splinter E, John S, Sabo PJ, et al. Diverse gene reprogramming events occur in the same spatial clusters of distal regulatory elements. Genome Res. 2011;21:697–706.
49. Le Dily F, Baù D, Pohl A, Vicent GP, Serra F, Soronellas D, et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. Genes Dev. 2014;28:2151–62.
50. Denissov S, van Driel M, Voit R, Hekkelman M, Hulsen T, Hernandez N, et al. Identification of novel functional TBP-binding sites and general factor repertoires. EMBO J. 2007;26:944–54.
51. Martens JHA, Brinkman AB, Simmer F, Francoijs K-J, Nebbioso A, Ferrara F, et al. PML-RARalpha/RXR alters the epigenetic landscape in acute promyelocytic leukemia. Cancer Cell. 2010;17:173–85.
52. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-seq (MACS). Genome Biol. 2008;9:R137.
53. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57.
54. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37:1–13.

Chapter 3

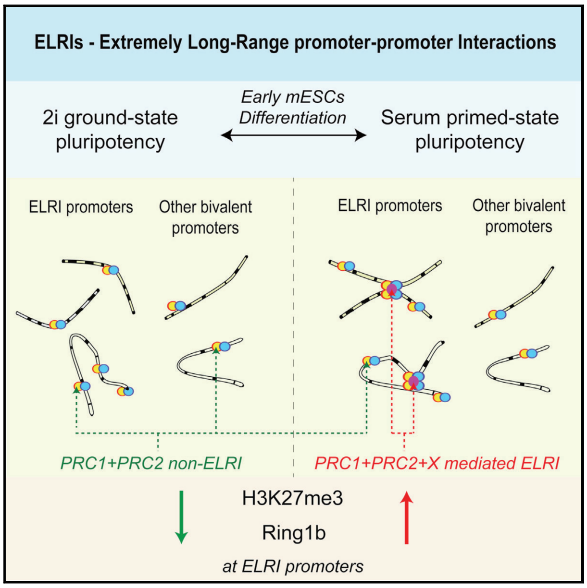
Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency

Short Article

Cell Stem Cell

Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency

Graphical Abstract



Highlights

- Capture Hi-C identified Extremely Long-Range Promoter-Promoter Interactions (ELRIs)
- ELRIs are established during the ground-state to primed pluripotency transition
- ELRIs involve *Hox* and a subset of bivalent loci
- H3K27me3 and Eed are necessary, but not sufficient, to mediate ELRIs

Authors

Onkar Joshi, Shuang-Yin Wang, Tatyana Kuznetsova, ..., Mikhail Spivakov, Daniel Burgess, Hendrik G. Stunnenberg

Correspondence

h.stunnenberg@ncmls.ru.nl

In Brief

Stunnenberg and colleagues used CHi-C to identify Extremely Long-Range Promoter-Promoter Interactions (ELRIs) in mESCs. Their analysis points to a spatiotemporal mechanism for repressing *Hox* and other developmentally important genes during the transition from the 2i ground-state to the primed serum state regulated by PRC2.

Accession Numbers

GSE72164

Joshi et al., 2015, Cell Stem Cell *17*, 748–757
December 3, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.stem.2015.11.010

CellPress

Cell Stem Cell
Short Article

Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency

Onkar Joshi,[1,5] Shuang-Yin Wang,[1,5] Tatyana Kuznetsova,[1] Yaser Atlasi,[1] Tianran Peng,[1] Pierre J. Fabre,[2] Ehsan Habibi,[1] Jani Shaik,[1] Sadia Saeed,[1,6] Lusy Handoko,[1,7] Todd Richmond,[3] Mikhail Spivakov,[4] Daniel Burgess,[3] and Hendrik G. Stunnenberg[1,*]

[1]Radboud University, Faculty of Science, Department of Molecular Biology, 6525GA Nijmegen, the Netherlands
[2]School of Life Sciences, Ecole Polytechnique Fédérale, Lausanne, 1015 Lausanne, Switzerland
[3]Roche NimbleGen, Incorporated, 500 South Rosa Road, Madison, WI 53719, USA
[4]Nuclear Dynamics Programme, Babraham Institute, Cambridge, Cambridgeshire CB22 3AT, UK
[5]Co-first author
[6]Present address: Department of Biochemistry, PMAS Arid Agriculture University Rawalpindi, 46300 Rawalpindi, Pakistan
[7]Present address: Division of Genomic Technologies, RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
*Correspondence: h.stunnenberg@ncmls.ru.nl
http://dx.doi.org/10.1016/j.stem.2015.11.010

SUMMARY

Serum-to-2i interconversion of mouse embryonic stem cells (mESCs) is a valuable in vitro model for early embryonic development. To assess whether 3D chromatin organization changes during this transition, we established Capture Hi-C with target-sequence enrichment of DNase I hypersensitive sites. We detected extremely long-range intra- and inter-chromosomal interactions between a small subset of H3K27me3 marked bivalent promoters involving the *Hox* clusters in serum-grown cells. Notably, these promoter-mediated interactions are not present in 2i ground-state pluripotent mESCs but appear upon their further development into primed-like serum mESCs. Reverting serum mESCs to ground-state 2i mESCs removes these promoter-promoter interactions in a spatiotemporal manner. H3K27me3, which is largely absent at bivalent promoters in ground-state 2i mESCs, is necessary, but not sufficient, to establish these interactions, as confirmed by Capture Hi-C on *Eed*$^{-/-}$ serum mESCs. Our results implicate H3K27me3 and PRC2 as critical players in chromatin alteration during priming of ESCs for differentiation.

INTRODUCTION

The pluripotent properties of mouse embryonic stem cells (mESCs) make them an invaluable model for fundamental research into the regulatory mechanisms in early development. mESCs are classically cultured in growth media supplemented with fetal calf serum and leukemia Inhibitory Factor (LIF) or, more recently, in serum-free 2i medium that contains LIF plus two small-molecule kinase inhibitors: PD0325901, targeting the mitogen-activated protein kinase (MEK) pathway, and CHIR99021, targeting the glycogen synthase kinase-3 (GSK3) pathway (Ying et al., 2008). It is well accepted that 2i mESCs represent a ground-state pluripotency, whereas the classical serum-derived mESCs are reminiscent of post-implantation pluripotent stem cells (Marks and Stunnenberg, 2014; Marks et al., 2012; Nichols and Smith, 2009; Odsworth et al., 2015; Plusa and Hadjantonakis, 2014; Ying et al., 2008).

mESCs grown in both serum+LIF ("serum mESCs") and 2i+LIF media ("2i mESCs") are pluripotent; however, they show distinct epigenetic landscapes and transcriptomic profiles (Habibi et al., 2013; Marks et al., 2012). 2i mESCs have higher expression levels of metabolic genes and diminished expression levels of lineage-priming genes compared to serum mESCs (Marks et al., 2012). In serum mESCs, around 3,000 genes have a bivalent chromatin state (co-occurrence of H3K27me3 and H3K4me3 marks) and are said to be poised for activation and predominantly coincide with promoters of genes involved in cell-fate determination and development (Azuara et al., 2006; Bernstein et al., 2006; Ku et al., 2008). In 2i mESCs, the number of bivalent genes reduces drastically; however, loss of repressive H3K27me3 does not result in transcriptional activation (Marks et al., 2012; Pasini et al., 2007). 2i mESCs have a hypo-methylated DNA similar to pre-implantation embryos, whereas serum mESCs are hyper-methylated, reminiscent of post-implantation embryos (Habibi et al., 2013; Leitch et al., 2013; Seisenberger et al., 2012; Smith et al., 2012). Together, these findings suggest that serum mESCs are epigenetically more restricted and developmentally primed as compared to ground-state 2i mESCs.

The 3D organization of chromatin may well play a role in initiation and/or maintenance of the distinct epigenetic landscapes and gene expression in the two states of pluripotency. The plasticity of the 3D conformation and its instructive role has been a subject of debate in recent years. The prevailing view is that in closely related cell types global interaction dynamics are limited, while fine-tuning of local interactions is potentially more frequent and linked to the transcriptional state (Bickmore, 2013; Dixon et al., 2015; Gibcus and Dekker, 2013; de Wit et al., 2013). Most of the promoter-promoter and promoter-enhancer

interactions take place within Topologically Associated Domains (TADs), which are on average 880 kb in size (Dixon et al., 2012). On the other hand, long-range interactions that span across TADs as well as chromosomes are rare when compared to the frequency of intra-TAD interactions (Gibcus and Dekker, 2013; Schwartz et al., 2012; Seitan et al., 2013; Sofueva et al., 2013).

In this study we assessed the dynamic reorganization of the 3D chromatin architecture in two closely related states of pluripotency using a Capture Hi-C (CHi-C) approach.

RESULTS

CHi-C

To study dynamics in chromatin architecture and to characterize long-range interactions, we performed Hi-C using DpnII as the restriction enzyme, potentially reaching a genome-wide coverage at a resolution of less than 1 kb. We subsequently performed enrichment of interactions by a target-capture (called CHi-C), similar to the exome sequencing approach. We enriched ligation products (interactions) using probes designed to capture targeted regions (Table S1). Probes were designed to capture loci targeting the union of all DNase I hyper-sensitive sites (DHSs) from the chromatin of serum and 2i mESCs (Figures S1A and S1B, Table S1). This elaborate representation of the genome of ~100k loci allows probing of promoter-promoter, promoter-enhancer, and enhancer-enhancer contacts (Figure 1A). After paired-end sequencing was performed, we removed unanchored reads and duplicates, yielding a total of 784 million reads (Table S1) with ~2 million significant long-range contacts. The fold enrichment for the probe-targeted open chromatin regions was ~20-fold with a capturing efficiency of 58.4% on average (at least one-end overlap with targeted regions). To validate the robustness of the CHi-C approach, all experiments were conducted in two biological replicates and resulted in high reproducibility (Figures S1C–S1H). The distributions of promoter-promoter, promoter-enhancer, and enhancer-enhancer interactions between 2i and serum mESCs were found to be very similar (Figure 1A).

Extremely Long-Range Interactions

We computed a score for each interaction that is a function of distance and frequency of contact between two genomic fragments, using a pipeline specifically developed to analyze CHi-C data, called "CHiCAGO" (Cairns et al., 2015). We used five di-tag reads within five adjacent DpnII fragments as a cut-off. Visual inspection showed high-ranking interactions involving *Hox* and other genomic loci in serum mESCs that were absent or very low in 2i CHi-C (Table S2). Next, we performed virtual-4C analyses to identify long-range interactions excluding intra-TAD interactions. We filtered for differential long-range interactions between serum and 2i mESCs (>3-fold change). The Circos plots revealed prominent interactions that are intra- as well as inter-chromosomal (Figure 1B, red lines), henceforth defined as Extremely Long-Range Interactions (ELRIs). Including *Hox* genes, 108 protein coding genes and 93 non-coding genes were identified to be directly overlapping with ELRI loci (Table S2). Most of the previously reported *Hox* interactions in serum mESCs (Vieux-Rochas et al., 2015) were independently identified by our CHi-C approach. ELRIs nearly exclusively occur in serum mESCs and are absent or strongly reduced in ground-state 2i mESCs. Local intra-TAD contacts appear largely unaffected (Figure 1B, blue lines, and Figure S1I). Examples of dynamic interactions between the *HoxD* locus and *Lmx1b* or *Wt1* loci on chromosome 2, as well as the reverse from *Lmx1b* to *HoxD* locus, *Lhx2*, and *Dlx1/Dlx2*, are illustrated (Figures 1C–1E). To validate the results obtained from CHi-C, we performed 4C on selected ELRI loci and found high consistency, showing the robustness of the CHi-C approach (Figure S2). FISH experiments proved to be challenging because of the refractory properties of 2i mESCs and did not yield confident results.

Characteristics of ELRI Loci

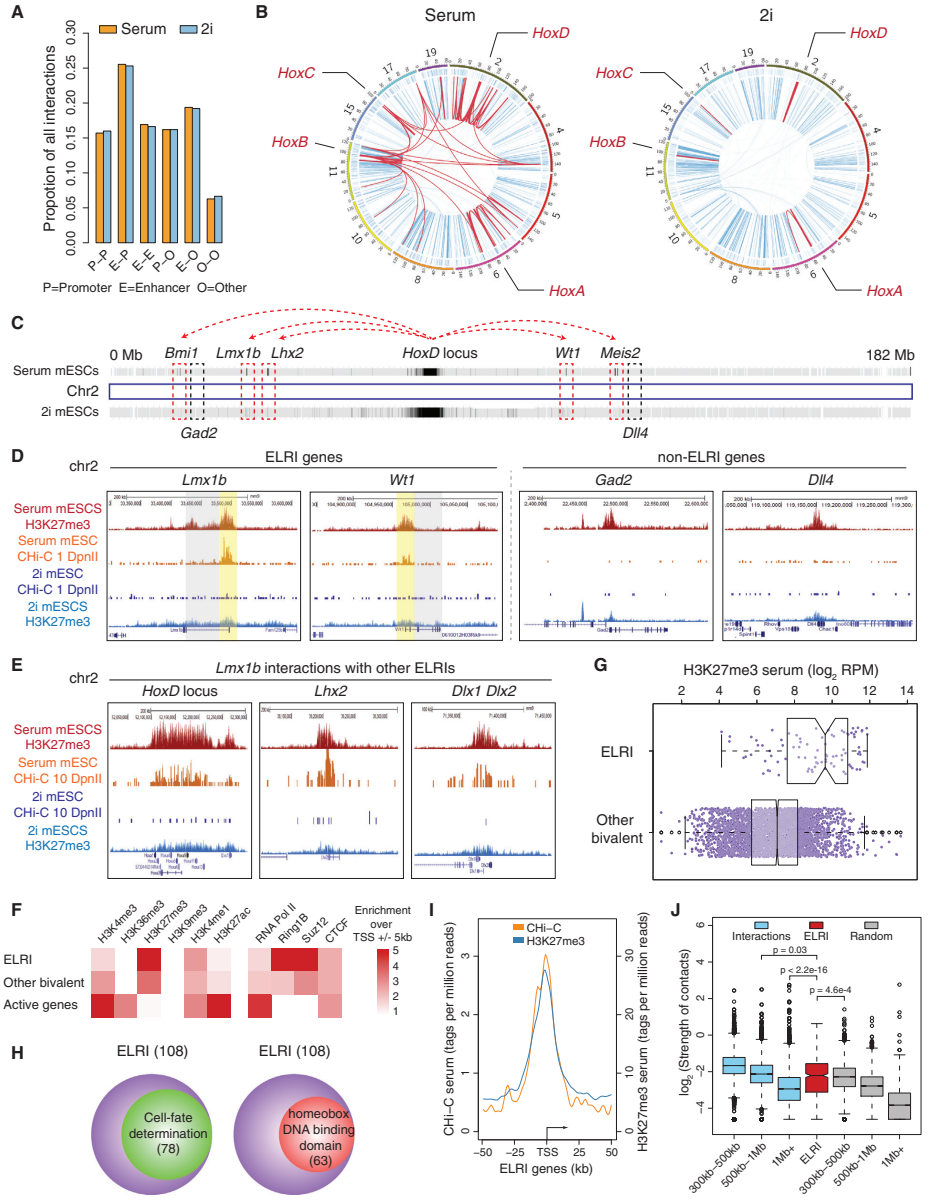
We next investigated the epigenetic makeup of ELRI loci by profiling histone marks: H3K4me3, H3K36me3, H3K4me1, H3K27ac, H3K9me3, and H3K27me3. In addition, we also profiled other DNA-binding proteins: RNA Polymerase II, Ring1B, Suz12, and CTCF (Figure 1F). The analysis reveals H3K27me3 as the prominent feature of ELRI loci. The presence of H3K27me3 with low levels of H3K4me3 is the hallmark of bivalent loci. Like H3K27me3, Ring1B and Suz12 are also prominent factors localized at ELRI loci (Figure 1F). The presence of H3K27me3 is not an exclusive feature of ELRI loci as the vast majority of bivalent loci, even those with very high H3K27me3 marking, do not participate in ELRIs (Figure 1G). A virtual-4C plot from the *HoxD* viewpoint underscores the selective nature of ELRIs (Figure 1D). Thus, the presence of H3K27me3 is a prominent, but not selective, feature of ELRIs.

It is well established that bivalent loci coincide with promoters of genes involved in cell-fate determination and development (Azuara et al., 2006; Bernstein et al., 2006). Accordingly, ELRI loci also largely overlap with promoters of transcription factors involved in cell-fate determination (78/108) and possess a homeobox DNA-binding domain (63/108) (Figure 1H, Table S2). Plotting the distributions of ELRI CHi-C and H3K27me3 tags reveals their spatial co-localization and confinement to a region (with an average size of 36 kb) centered on the transcription start sites (Figure 1I). The intimate connection between ELRI contacts and H3K27me3 can also be appreciated from *Lmx1b* and *Wt1* ELRI promoters (Figures 1C and 1D).

After establishing the connection between ELRI loci and co-localization with promoters, we compared the strength of ELRIs with other classes of promoter-promoter interactions. We computed the average normalized di-tags for promoter-promoter interactions representing intra-TAD interactions (300 kb to 1 mb), inter-TAD interactions (>1 mb), and ELRI contacts identified by the CHiCAGO pipeline (Figure 1J). Notwithstanding the extreme long distance, ELRI contacts were similar in strength to promoter-promoter intra-TAD contacts and were stronger than non-ELRI promoter-promoter inter-TAD interactions (Figure 1J).

Loss of ELRIs in *Eed*$^{-/-}$ Serum mESCs

Epigenetic profiling of ELRI promoters points to a role of H3K27me3 in the interactions. To investigate the putative role of PRC2 and H3K27me3 deposition in ELRIs, we performed CHi-C on *Eed*$^{-/-}$ mESCs cultured in serum media. Eed is one of the core components of the PRC2 complex and the absence of Eed results in destabilization of the PRC2 complex and a complete loss of H3K27me3 (Boyer et al., 2006; Wang et al., 2002). We observed that in *Eed*$^{-/-}$ mESCs, ELRIs are lost or strongly



(legend on next page)

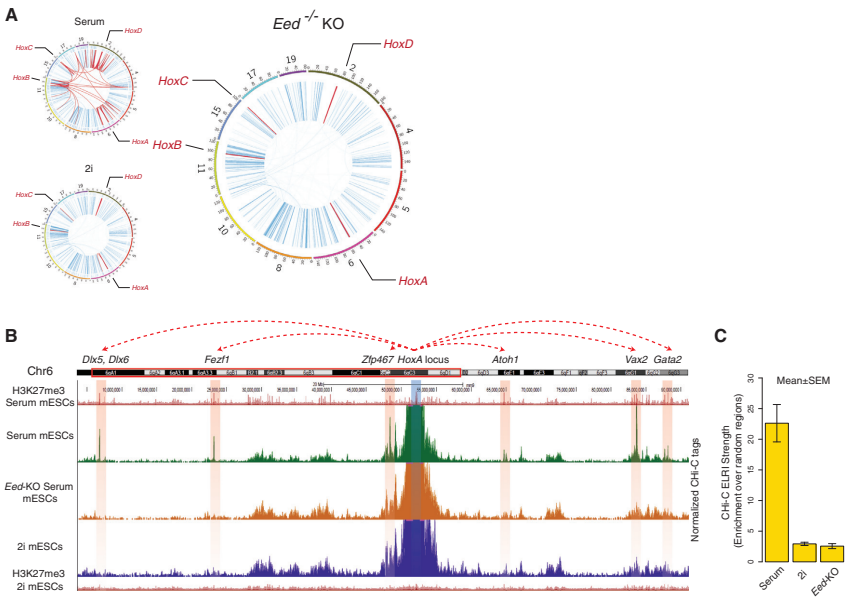


Figure 2. *Eed*$^{-/-}$ mESCs Lack ELRI Contacts as in 2i mESCs

(A) Circos plots for serum and 2i mESCs (left two plots) and for *Eed*$^{-/-}$ (right). Description is as in Figure 1B.

(B) An example showing the lack of ELRIs in *Eed*$^{-/-}$ mESCs as in 2i mESCs. Using the *HoxA* locus (blue shaded regions) as the viewpoint, CHi-C signals on chromosome 6 (sum of normalized di-tags in a sliding window of 100 DpnII fragments) are plotted for serum, 2i, and *Eed*$^{-/-}$ mESCs. Differential ELRI regions are shaded in red. On top, ELRI contacts are indicated by red arches.

(C) Bar plots of relative strength of ELRI contacts in serum, 2i, and *Eed*$^{-/-}$ mESCs. Standard error of the mean (SEM) of the enrichment is calculated using the two biological replicates.

Also see Figure S2.

reduced, similar to the situation in 2i mESCs (Figures 2A and 2B). To quantify the loss of ELRI strength, we compared average normalized di-tags on ELRI promoters of *Eed*$^{-/-}$ mESCs with WT serum mESCs and 2i mESCs (Figure 2C). The strength of ELRIs in *Eed*$^{-/-}$ mESCs was comparable to the strength observed in 2i mESCs. In contrast, the intra-TAD interactions (Figure 2A, blue

Figure 1. ELRIs in Serum mESCs are Lost in 2i mESCs

(A) The bar plots show the abundance of different classes of all CHi-C interactions in serum and 2i mESCs.

(B) A graphical representation of ELRIs using a Circos plot for the chromosomes involved (outermost ring). The plots show the interactions for all H3K27me3 marked loci (blue outer ring) in serum mESCs (left) and 2i (right). Inter- and intra-chromosomal ELRI contacts (red lines) and interactions between other bivalent genes (blue lines) are represented in the inner plots. The four *Hox* gene clusters are marked in red.

(C) Schematic representation of chromosome 2 and positions of genes, indicated by dashed squares. Density plots of CHi-C signals are plotted along with the chromosome. Two ELRI contacts (*Lmx1b* and *Wt1*) interacting with *HoxD* are highlighted in red, while non-ELRI genes (*Dll4* and *Gad2*) are highlighted in black.

(D) Zoom-in snapshots of browser views for *Lmx1b*, *Wt1*, *Dll4*, and *Gad2* from the *HoxD* locus as viewpoints. Snapshots for *Lmx1b* and *Wt1* depict differential ELRI contacts with the *HoxD* locus, in serum mESCs (orange track) and 2i mESCs (dark blue track). Single DpnII fragment resolution of CHi-C shows ELRI contacts focalized on promoters that overlap with H3K27me3 peaks in serum mESCs and are absent in 2i mESCs. H3K27me3 tracks for serum mESCs and 2i mESCs are colored red and light blue, respectively.

(E) Zoom-in snapshots as in (D) from *Lmx1b* as the viewpoint, showing interaction patterns with the *HoxD* locus, *Lhx2*, and *Dlx1*/*Dlx2*.

(F) A heat map showing topography of histone marks and DNA-binding factors on ELRI loci compared to other bivalent and active genes. Color scheme indicates highest to lowest enrichment.

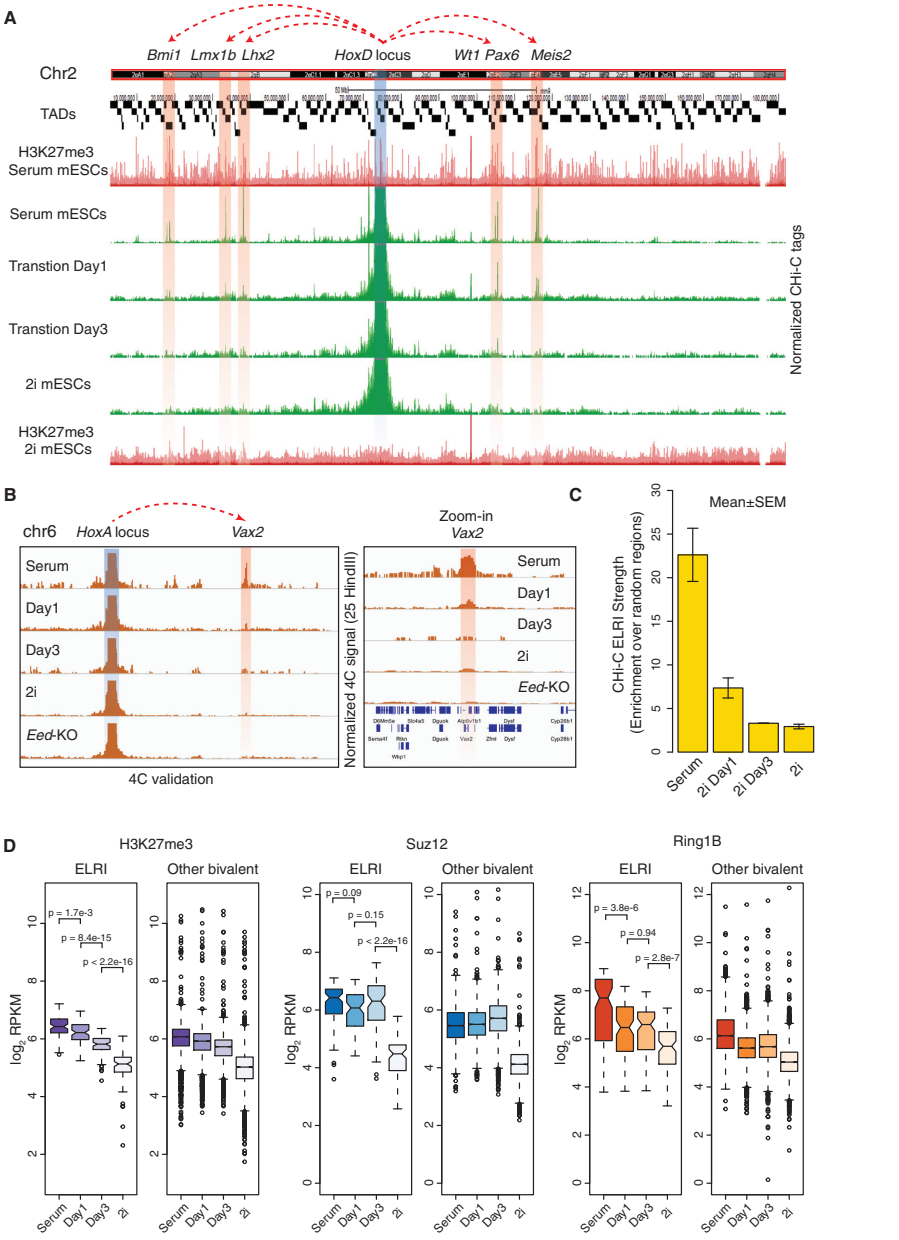
(G) A box plot showing comparable H3K27me3 intensity (log2RPKM) on ELRI and other bivalent loci.

(H) Diagrams of the gene-functional classes of ELRIs.

(I) Average profiles of CHi-C and H3K27me3 signals from serum mESCs on the transcription start sites (TSSs) of ELRI genes. CHi-C signal (orange line) shows the focal nature of ELRI contacts around the TSSs of ELRI genes and its correlation with the H3K27me3 mark (blue line).

(J) Comparison between strengths of ELRIs and all other promoter-promoter contacts. Promoter-promoter contacts are illustrated in three distance intervals: 300 kb to 500 kb, 500 kb to 1 mb, and >1 mb. The strengths of other significant promoter-promoter interactions, ELRI contacts, and random promoter-promoter contacts are represented in blue, red, and gray, respectively. The definition of strength can be accessed in the Supplemental Information.

Also see Figure S1.



(legend on next page)

line) were generally not affected by the loss of the H3K27me3 mark in *Eed*$^{-/-}$ mESCs as in 2i mESCs (Figure 2A). This implicates that unlike ELRIs, these intra-TAD interactions are not dependent on H3K27me3 and PRC2. We validated the *Eed*$^{-/-}$ CHi-C experiments by 4C on *HoxA* and *HoxD* as well as on selected ELRI promoters (Figures S2A and S2B). We also compared *Eed*$^{-/-}$ CHi-C data with existing 4C data, showing good overlap (Figure S2C, Denholtz et al., 2013). Thus, PRC2 and H3K27me3 play indispensable roles in the formation of ELRIs.

Time-Dependent Loss of ELRIs during Transition from Serum-to-2i mESCs

To study the dynamics of ELRIs during transition from serum-to-2i, we performed CHi-C on Day1 and Day3 after medium exchange (Figure 3A). The interaction patterns of ELRIs show a gradual loss starting as early as Day1. Most ELRI contacts are drastically reduced or were not detected anymore at Day3, such as the interaction between *Bmi1* and the *HoxD* locus (Figure 3A). Other ELRI contacts such as *Meis2* and *Pax6* are still observable at Day3, although at severely reduced strengths (Figure 3A). This dynamic loss of ELRI strength was validated using the 4C approach on Day1 and Day3 mESCs, as exemplified by *HoxA* and *Vax2* (Figures 3B and S3A) as well as other ELRI contacts (Figure S3B). To quantify the loss of ELRI strength we compared average normalized di-tags on ELRI promoters during the transition of serum-to-2i mESCs (Figure 3C). Compared to the strength of ELRIs in WT serum mESCs, we observed on average a 67% reduction on Day1 and an 85% reduction on Day3.

The loss of ELRIs in *Eed*$^{-/-}$ and in WT 2i mESCs shows that H3K27me3 deposited by the PRC2 complex is necessary for ELRIs. However, H3K27me3 deposition is not sufficient for ELRIs, because most bivalent loci that are equally marked with H3K27me3 are not involved in ELRIs. Given the interplay between PRC2 and PRC1 in epigenetic shaping of repressive chromatin, we investigated the role of PRC1 in ELRIs as PRC1 has been shown to be recruited to H3K27me3 to reinforce the function of PRC2 (Ku et al., 2008). We performed ChIP-seq experiments on Suz12 and Ring1B during the serum-to-2i transition. The occupancy of the PRC2 component Suz12 on ELRI loci was unaltered in serum mESCs even up to Day3 (Figure 3D, mid-panel), whereas H3K27me3 and ELRI contacts were largely lost (Figure 3D, left panel). In contrast, Ring1B, a central component of PRC1, shows a gradual reduction starting as early as Day1, and the loss of occupancy becomes pronounced or completed at Day3 in excellent synchrony with the loss of ELRIs (Figure 3D, right panel). The reduced occupancy observed is not due to reduced expression of Ring1B or other detected PRC components as determined by quantitative Mass Spectrometry (Figure S3C).

This temporal concordance between loss of ELRIs and PRC1 indicates a role of PRC1 as a reader of H3K27me3 and possibly in mediating ELRIs. We observed that bivalent genes that do not display ELRIs also have occupancy of PRC1 (Ring1B) similar to that of ELRI loci. During the serum-to-2i transition, bivalent loci in general show the same kinetics as ELRI loci with respect to loss of Ring1B (Figure 3D). It seems likely that ELRIs require a specific composition of the modular PRC1, specific transcription factor co-binding, epigenetic context, or some combination thereof that acts only on ELRI loci.

ELRI-Related Local Chromatin Dynamics and Gene Transcription

We next investigated the effect of ELRIs on gene expression. We previously showed that in 2i, two-thirds of the bivalent genes (total ~3,000) are not yet marked with the H3K27me3, but only ~10% show significant transcriptional activity (Marks et al., 2012). Similarly, ~18% of ELRI genes are significantly expressed in 2i, and they become repressed in serum having gained H3K27me3 and ELRIs (Figures 4A–4B, Table S3). Strikingly, these genes encode subunits of PRC1: Cbx4, Cbx8, and Bmi1 (Morey et al., 2012). This suggests a potential change in the composition of PRC1 in the transition from ground-state 2i to primed serum mESCs. Another ELRI gene that is significantly more highly expressed in 2i mESCs is *Tbx3*, which has been shown to be important for self-renewal of mESCs (Ivanova et al., 2006).

Next, we looked at the temporal changes in local chromatin architecture of the ELRI genes during the transition from the serum-to-2i mESCs state. Loss of ELRIs coincides with diverse spatiotemporal changes in local intra-TAD interaction patterns and gene expression. The diverse changes in the local 3D chromatin structure and their effects on gene expression are illustrated by the following four examples.

ELRI loci such as those encoding Bmi1 gradually lost the H3K27me3 mark, and their interactions with *Hox* and with local inter-TAD H3K27me3 marked regions were also lost. These changes were accompanied by gain of H3K27ac at the *Bmi1* promoter, i.e., a change in chromatin state (Figure 4C, orange shaded region), and they resulted in increased transcriptional activity. The *Lbx1* locus also gradually lost "negative" H3K27me3-mediated interactions, while some of the interacting regions in serum mESCs, which had been decorated with the repressive H3K27me3, gained "positive" histone modifications H3K27ac and H3K4me1 in 2i mESCs (Figure 4F, green shaded region). In contrast to *Bmi1*, the transcriptional activity of *Lbx1*

Figure 3. Dynamics of ELRIs and PRC Subunit Occupancy during the Serum-to-2i Transition

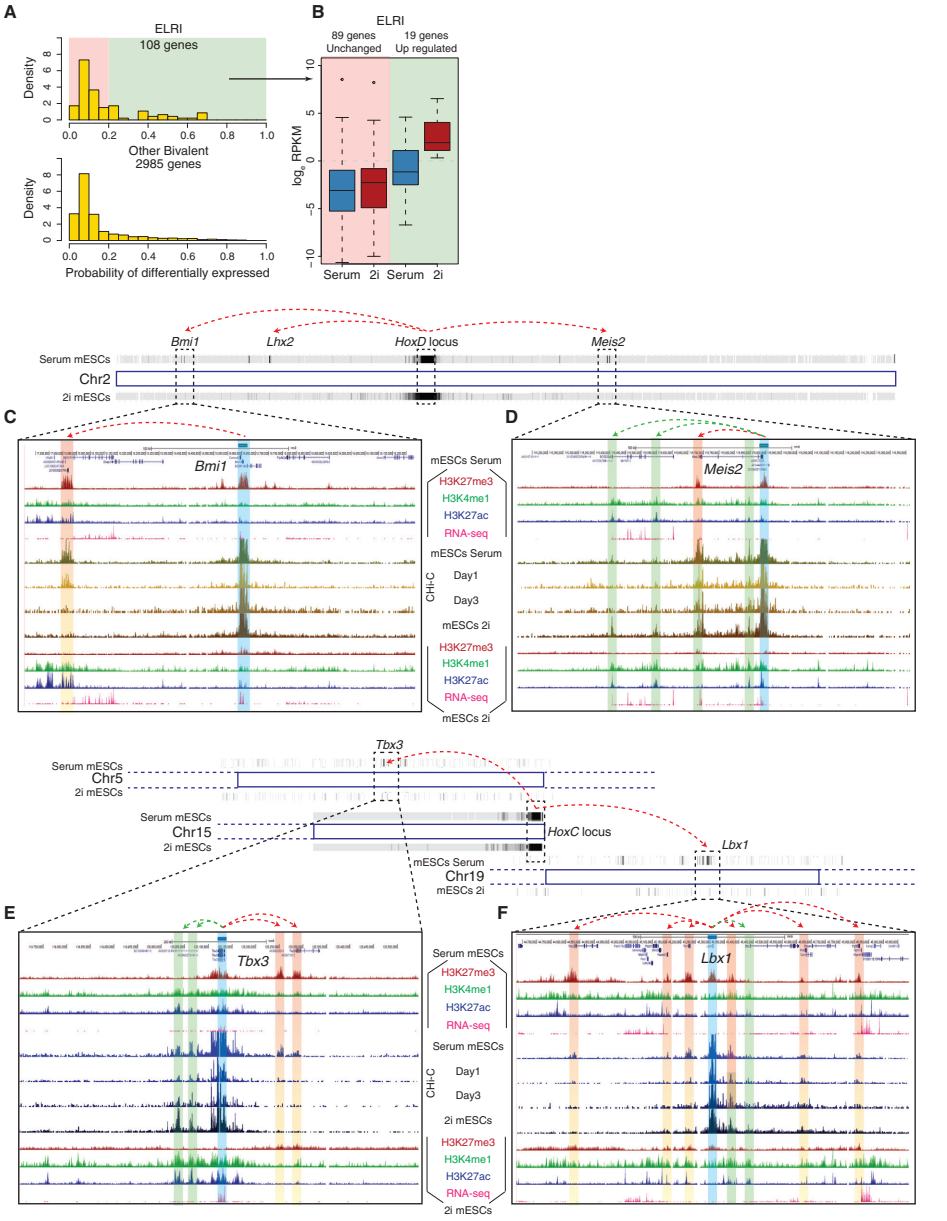
(A) Differential ELRI contacts during the serum-to-2i transition. Using the *HoxD* locus (blue shaded regions) as the viewpoint, CHi-C signals on chromosome 2 (sum of normalized di-tags in a sliding window of 10 DpnII fragments) are plotted for serum, Day1, and Day3 of serum-to-2i transition and 2i. Dynamics of ELRIs are highlighted as red shaded regions. On top, ELRI contacts are indicated by red arches.

(B) A browser view of interaction between the viewpoint *HoxA* locus (shaded blue) and *Vax2* (shaded red) during the serum-to-2i transition and in *Eed*$^{-/-}$ mESCs (sliding window of 25 HindIII fragments). Red dashed arch indicates ELRI contact. A zoom-in view on the right shows the loss of ELRI on *Vax2* during the serum-to-2i transition and in *Eed*$^{-/-}$ mESCs.

(C) Bar plots of relative strength of ELRI contacts in serum, Day1, and Day3 of serum-to-2i transition and 2i. Standard error of the mean (SEM) of the enrichment is calculated using the two biological replicates.

(D) Box plots for intensities of H3K27me3, Suz12, and Ring1B on ELRI loci and other bivalent genes in serum, Day1, and Day3 of serum-to-2i transition and 2i mESCs.

Also see Figure S3.



(legend on next page)

was not significantly increased (Figure 4F). In the case of *Tbx3*, interactions with H3K27me3 loci were lost (Figure 4E, orange shaded regions), and contacts with enhancer-like loci (marked with H3K27ac and H3K4me1) were strengthened (Figure 4E, green shaded region). However, unlike *Lbx1*, the transcriptional activity was increased in response to local chromatin changes. Finally, at the *Meis2* locus, the loss of ELRI and H3K27me3 resulted in the formation of novel contacts between the promoter and potential enhancers (H3K27ac and H3K4me1 marked loci) and enhanced transcription of *Meis2* (Figure 4D, green shaded regions).

ELRI genes move from a relatively active chromatin state (in 2i mESCs) to a repressed chromatin state (in serum mESCs) when these changes are considered in a developmental context (2i-to-serum state pluripotency). The changes in the chromatin state are illustrated in a heat map (Figure S4A). The gradual gain of H3K27me3 is evident on all ELRI loci during the transition from ground-state to primed mESCs. However, only a subset of ~30 genes lose H3K27ac during the transition from the 2i-to-serum state and are transcriptionally repressed.

DISCUSSION

In this study, we show that ELRIs involving *Hox* and other genomic loci are present in serum mESCs but not yet established in the ground-state 2i condition. Our analysis provides evidence for spatiotemporal changes in 3D chromatin structure involving establishment of ELRIs during the 2i-to-serum transition. The overall organization of ELRI contacts established during this transition suggests a coordinated program that encompasses extensive reshaping of the transcriptome, epigenome, and 3D interactome during early stem cell differentiation. The absence of ELRIs in 2i ground-state seems to indicate that these interactions are not essential for the maintenance of pluripotency. Based on the idea that 2i mESCs represent an earlier developmental state (i.e. the inner cell mass [ICM]) and serum, a later developmental state (Boroviak et al., 2014; Habibi et al., 2013; Hackett and Surani, 2014; Huang et al., 2014; Marks and Stunnenberg, 2014; Nichols and Smith, 2009; Odsworth et al., 2015; Plusa and Hadjantonakis, 2014; Ying et al., 2008), we hypothesize that ELRIs are absent in embryos prior to implantation and that these interactions are acquired at later stages, probably to restrict or poise controlled genes for transcriptional activity. This proposal would be in line with the previous findings that H3K27me3 and PRC2 are not essential for pre-implantation embryos, but are essential for differentiation and formation of primary cell layers (Pasini et al., 2007; Wang et al., 2002).

The severe reduction in levels of H3K27me3 at ELRI loci in 2i mESCs, compared to serum mESCs, hints at a role of PRCs in ELRIs. CHi-C on *Eed*$^{-/-}$ showed that PRC2 activity is necessary. The occupancy of ELRI loci by Suz12 remains unaltered up to Day3, at which point 85% of ELRI strength is lost, showing that the enzymatic activity of PRC2 is a critical factor, rather than its binding at ELRI loci.

The synchrony between loss of ELRIs and loss of the PRC1 component Ring1B is suggestive of its role in ELRIs. The traditional view is that PRC2 acts as the initiator in the formation of a PRC2/PRC1 chromatin state by depositing H3K27me3 that recruits PRC1 through the Cbx reader protein family (Boyer et al., 2006; Simon and Kingston, 2009). In serum mESCs, PRC2 has been shown to recruit PRC1 components to *Hox* loci (Boyer et al., 2006). Hence, it is possible that PRC2 acts only as the recruiter of PRC1, which then mediates ELRIs. In line with this model, knocking out *Eed* precludes H3K27me3 deposition and consequently recruitment of a PRC1 complex (Boyer et al., 2006; Wang et al., 2002). In fully converted 2i mESCs (Day 15–18), similar to *Eed*$^{-/-}$ mESCs, the H3K27me3 mark is strongly diminished, resulting in poor recruitment of PRC1 (Habibi et al., 2013). We postulate that, in 2i mESCs, the loss of enzymatic activity of Ezh2 results in poor or lack of H3K27me3 deposition; consequently, canonical PRC1 cannot bind and ELRIs cannot be established. In the past, Denholtz et al. (2013) have also suggested a role of PRC2 in chromatin compaction in serum mESCs, although the HindIII-based 4C did not have sufficient resolution to reveal the promoter-promoter nature of ELRI contacts described in this study.

The subunit composition of the PRC1 complex involved in ELRIs remains to be established. Our data show that Ring1B occupancy is not only reduced at ELRI loci but is also reduced at other bivalent loci and displays the same kinetics during the transition. Thus, the mere presence of Ring1B (PRC1) is insufficient to generate selectivity in ELRIs. A role of Ring1B in canonical PRC1 has been reported in local compaction at *Hox* loci (Eskeland et al., 2011). Furthermore, Eskeland and coworkers using the *Eed*$^{-/-}$ mESCs, showed that PRC2 is not sufficient for local compaction of chromatin in mESCs, providing a striking parallel to our study. Given the highly modular composition of PRC1 complexes (Chen and Dent, 2014; Creppe et al., 2014; Senthilkumar and Mishra, 2009), we hypothesize that a specific PRC1 subunit composition is required for ELRI formation, and this specific PRC1 complex distinguishes ELRI loci from other bivalent loci.

While our manuscript was in revision, Schoenfelder and coworkers identified extremely long-range interactions in serum mESCs using a promoter CHi-C approach. Based on *Ring1A-Ring1B*-dKO mESCs, they pointed to a role of PRC1, in line with our speculation (Schoenfelder et al., 2015). Their promoter and our DHS capture approaches identify ELRIs that largely overlap (80% of coding genes in Schoenfelder et al., 2015 are independently identified in our study as ELRIs; Figure S4B). While we firmly established a role of PRC2 using CHi-C on the *Eed*$^{-/-}$ mESCs, they revealed the critical role of PRC1 in the

Figure 4. Transcriptional and Chromatin State Changes at ELRI Loci

(A and B) Left panels: a histogram for probabilities of differential expression for 108 ELRI genes (top) and 2,985 other bivalent genes (bottom); genes with a probability value of greater than 0.2 were considered as differentially expressed (represented as green). Right panel: of the 108 ELRI genes, 89 genes remained unchanged in transcriptional activity and 19 genes were upregulated in 2i mESCs compared to the serum condition.

(C–F) Snapshots of ELRI loci of *Bmi1*, *Meis2*, *Tbx3*, and *Lbx1* genes, respectively. ELRI genes (blue shaded) interact with local H3K27me3 regions (orange shaded). While resolving H3K27me3-marked "negative" interactions, new "positive" interactions between the promoters and H3K27ac-marked regions (green shaded) are established. On top of the black boxes for the exemplified loci, intra- and inter-chromosomal interactions between *Hox* loci and other ELRI loci are represented as density plots. The red dashed arches represent repressive interactions, whereas the green dashed arches represent positive interactions.

Also see Figure S4.

process. In the *Ring1A-Ring1B*-dKO mESCs, ELRIs cannot be established because these dKO mESCs have an active PRC2 complex but a defunct PRC1 complex. Taking both studies together, it is prudent to postulate that PRC2 acts as an initiator of ELRIs by deposition of H3K27me3 and subsequent recruitment of canonical PRC1, which may act as the physical mediator of ELRIs.

EXPERIMENTAL PROCEDURES

Cell Culture

E14 Tg2a (also called E14) is a male mESC line of 129/Ola background. Serum (E14) mESCs were grown in DMEM containing 10% fetal calf serum in the presence of LIF, referred to as serum medium. 2i (E14) mESCs were grown in serum-free NDiff 227 supplemented with MEK inhibitor PD0325901 (1μM) and GSK3 inhibitor CHIR99021 (3μM) in the presence of LIF, referred to as 2i medium (Ying et al., 2008). All cell cultures were conducted in feeder-free conditions. The serum-to-2i transition was carried out by washing the mESCs in serum medium twice with PBS and then switching to 2i medium. *Eed*$^{-/-}$ mESCs were provided to us by Luciano di Croce and Anton Wutz, and are described in Morey et al. (2012) and Schoeftner et al. (2006). *Eed*$^{-/-}$ mESCs were also grown in identical conditions to those of WT E14 serum mESCs.

CHi-C

The CHi-C experiment was divided into two parts: in-nucleus Hi-C and ssDNA probe capture enrichment. In-nucleus Hi-C was carried out as described in Nagano et al. (2015). DpnII was used as the restriction enzyme. On beads DNA amplification PCR was carried out with seven to nine cycles to generate around 1 μg of Hi-C library DNA. The ssDNA probe capture step was carried out using the protocol provided by Roche NimbleGen Inc. (http://sequencing.roche.com/products/nimblegen-seqcap-target-enrichment/seqcap-ez-system/seqcap-ez-developer.html) optimized for the probe capture library. Libraries were indexed using NEXTflex adapters (Bioo Scientific Corporation) and 75 bp or 43 bp paired-end sequencing was performed on Illumina instruments using TruSeq reagents (Illumina), according to the manufacturer's instructions.

DNase I-Seq, ChIP-Seq, 4C, and RNA-Seq

A detailed description of sample preparation and data analysis is available in the Supplemental Experimental Procedures.

Public Datasets Used in This Study

4C primers and 4C data for *Eed*$^{-/-}$ mESCs and serum mESCs are from Denholtz et al. (2013). The CHi-C *Hox* loci interactors list is from Schoenfelder et al. (2015). H3K4me3, H3K9me3, and H3K36me3 ChIP-seq data are from Marks et al. (2012).

ACCESSION NUMBERS

The accession number for all the raw data generated using high-throughput sequencing reported in this paper has been deposited in the GEO database (GEO: GSE72164).

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes four figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2015.11.010.

AUTHOR CONTRIBUTIONS

O.J. performed CHi-C experiments, ChIPs, 4C experiments, visualization, and writing. S.-Y.W. performed bioinformatics analysis, visualization, and writing. T.K. performed 4C validation experiments. Y.A. performed ChIP experiments and RNA-seq. T.P. helped with ChIP-seq. E.H. and J.S. performed proteomic studies. S.S. performed DNase I-seq. L.H. performed ChIPs. T.R. and D.B. designed and manufactured the probes. M.S. provided assistance with the CHiCAGO analysis pipeline. P.J.F. provided experimental expertise and advice. H.G.S. is responsible for concepts, supervision, and the writing of the manuscript.

ACKNOWLEDGMENTS

We thank members of the stem cell group and Hendrik Marks in particular for discussions, and Luciano di Croce and Anton Wutz for the *Eed*$^{-/-}$ KO mESCs. We thank Eva Janssen-Megens, Kim Berentsen, Nilofar Sharifi, and Bowon Kim for sequencing; Jonathan Cairns and Paula Freire Pritchett for help with the CHiCAGO pipeline; Denise Raterman, Jennifer Wendt, and Heidi Rosenbaum for advice with the capture part of the CHi-C technique; Denis Duboule for advice on the subject; and Biola M. Javierre for advice on the traditional Hi-C protocol. The computational work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. This work was supported by ERC grant ERC-2013-AdG No. 339431 – SysStemCell.

Received: August 13, 2015
Revised: October 11, 2015
Accepted: November 13, 2015
Published: December 3, 2015

REFERENCES

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., and Fisher, A.G. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–538.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Bickmore, W.A. (2013). The spatial organization of the human genome. Annu. Rev. Genomics Hum. Genet. *14*, 67–84.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat. Cell Biol. *16*, 516–528.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349–353.

Cairns, J., Freire-Pritchett, P., Wingett, S.W., Dimond, A., Plagnol, V., Zerbino, D., Schoenfelder, S., Javierre, B.-M., Osborne, C., Fraser, P., et al. (2015). CHiCAGO: Robust Detection of DNA Looping Interactions in Capture Hi-C data. bioRxiv. http://dx.doi.org/10.1101/028068.

Chen, T., and Dent, S.Y.R. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nat. Rev. Genet. *15*, 93–106.

Creppe, C., Palau, A., Malinverni, R., Valero, V., and Buschbeck, M. (2014). A Cbx8-containing polycomb complex facilitates the transition to gene activation during ES cell differentiation. PLoS Genet. *10*, e1004851.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227–231.

Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., Pellegrini, M., and Plath, K. (2013). Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. Cell Stem Cell *13*, 602–616.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A.I., Wutz, A., and Bickmore, W.A. (2011). Ring1B Compacts

Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. Mol. Cell *38*, 452–464.

Gibcus, J.H., and Dekker, J. (2013). The hierarchy of the 3D genome. Mol. Cell *49*, 773–782.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell *13*, 360–369.

Hackett, J.A., and Surani, M.A. (2014). Regulatory principles of pluripotency: from the ground state up. Cell Stem Cell *15*, 416–430.

Huang, K., Maruyama, T., and Fan, G. (2014). The naive state of human pluripotent stem cells: a synthesis of stem cell and preimplantation embryo transcriptome analyses. Cell Stem Cell *15*, 410–415.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533–538.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., et al. (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. *4*, e1000242.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat. Struct. Mol. Biol. *20*, 311–316.

Marks, H., and Stunnenberg, H.G. (2014). Transcription regulation and chromatin structure in the pluripotent ground state. Biochim. Biophys. Acta *1839*, 129–137.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., and Stunnenberg, H.G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590–604.

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S.A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell *10*, 47–62.

Nagano, T., Várnai, C., Schoenfelder, S., Javierre, B.-M., Wingett, S.W., and Fraser, P. (2015). Comparison of Hi-C results using in-solution versus in-nucleus ligation. Genome Biol. *16*, 175.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Odsworth, B.E.T.D., Lynn, R.O.F., and Owley, S.A.A.C. (2015). The Current State of Naive Human Pluripotency. Stem Cells *33*, 3181–3186.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol. Cell. Biol. *27*, 3769–3779.

Plusa, B., and Hadjantonakis, A.-K. (2014). Embryonic stem cell identity grounded in the embryo. Nat. Cell Biol. *16*, 502–504.

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. EMBO J. *25*, 3110–3122.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., et al. (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat. Genet. *47*, 1179–1186.

Schwartz, Y.B., Linder-Basso, D., Kharchenko, P.V., Tolstorukov, M.Y., Kim, M., Li, H.B., Gorchakov, A.A., Minoda, A., Shanower, G., Alekseyenko, A.A., et al. (2012). Nature and function of insulator protein binding sites in the Drosophila genome. Genome Res. *22*, 2188–2198.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell *48*, 849–862.

Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., et al. (2013). Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. Genome Res. *23*, 2066–2077.

Senthilkumar, R., and Mishra, R.K. (2009). Novel motifs distinguish multiple homologues of Polycomb in vertebrates: expansion and diversification of the epigenetic toolkit. BMC Genomics *10*, 549.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nat. Rev. Mol. Cell Biol. *10*, 697–708.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339–344.

Sofueva, S., Yaffe, E., Chan, W.-C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J. *32*, 3119–3129.

Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. Proc. Natl. Acad. Sci. USA *112*, 4672–4677.

Wang, J., Mager, J., Schnedier, E., and Magnuson, T. (2002). The mouse PcG gene eed is required for Hox gene repression and extraembryonic development. Mamm. Genome *13*, 493–503.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Cell Stem Cell

Supplemental Information

Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency

Onkar Joshi, Shuang-Yin Wang, Tatyana Kuznetsova, Yaser Atlasi, Tianran Peng, Pierre J. Fabre, Ehsan Habibi, Jani Shaik, Sadia Saeed, Lusy Handoko, Todd Richmond, Mikhail Spivakov, Daniel Burgess, and Hendrik G. Stunnenberg

Figure S1

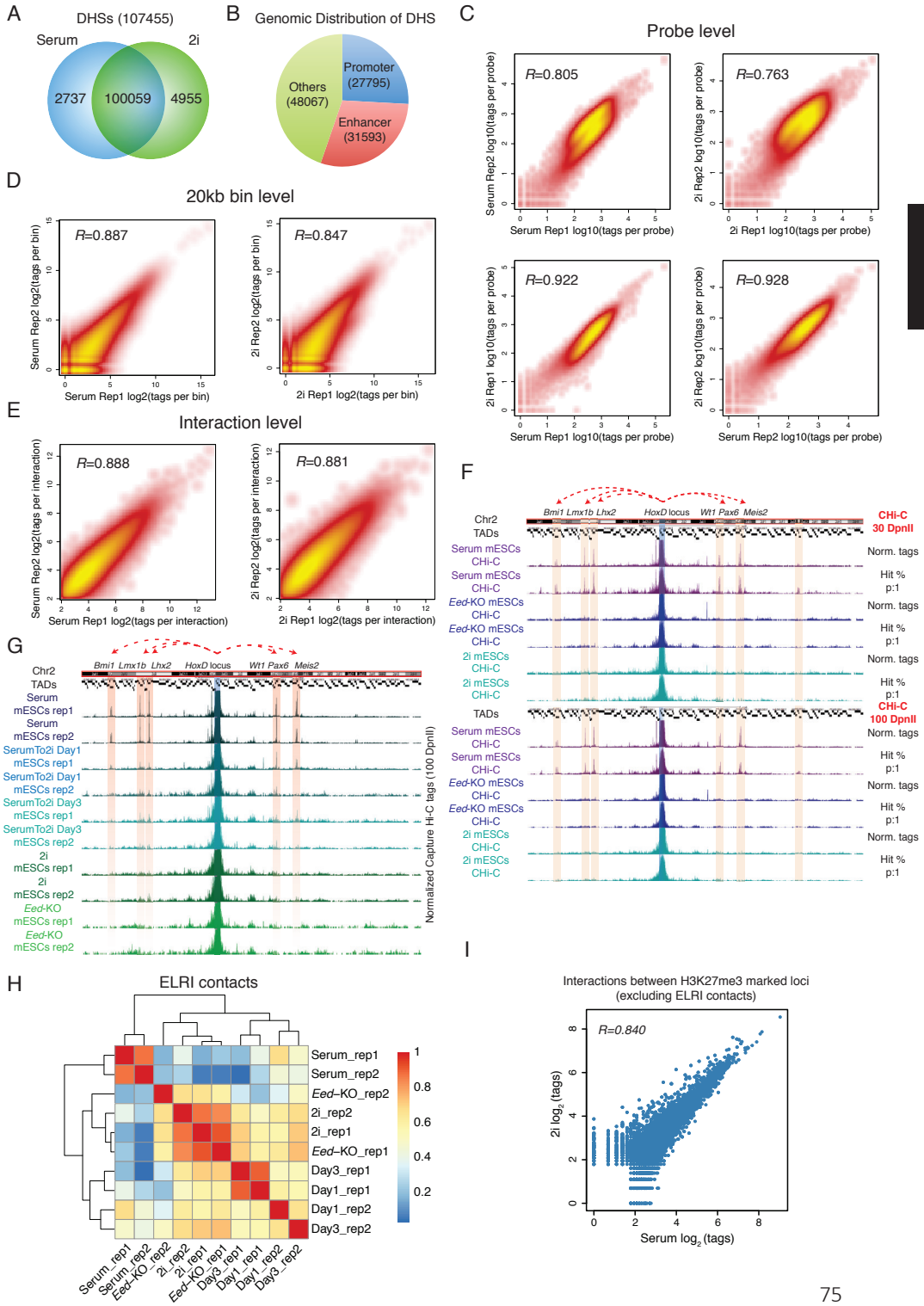

A
DHSs (107455)
Serum
2i
2737
100059
4955
B
Genomic Distribution of DHS
Promoter (27795)
Enhancer (31593)
Others (48067)
C
Probe level
R=0.805
R=0.763
R=0.922
R=0.928
Serum Rep2 log10(tags per probe)
Serum Rep1 log10(tags per probe)
2i Rep2 log10(tags per probe)
2i Rep1 log10(tags per probe)
D
20kb bin level
R=0.887
R=0.847
Serum Rep2 log2(tags per bin)
Serum Rep1 log2(tags per bin)
2i Rep2 log2(tags per bin)
2i Rep1 log2(tags per bin)
E
Interaction level
R=0.888
R=0.881
Serum Rep2 log2(tags per interaction)
Serum Rep1 log2(tags per interaction)
2i Rep2 log2(tags per interaction)
2i Rep1 log2(tags per interaction)
F
Bmi1 Lmx1b Lhx2
HoxD locus
Wt1 Pax6 Meis2
Chr2
TADs
CHi-C 30 DpnII
CHi-C 100 DpnII
Serum mESCs CHi-C
Eed-KO mESCs CHi-C
2i mESCs CHi-C
Norm. tags
Hit % p:1
G
Serum mESCs rep1
Serum mESCs rep2
SerumTo2i Day1 mESCs rep1
SerumTo2i Day1 mESCs rep2
SerumTo2i Day3 mESCs rep1
SerumTo2i Day3 mESCs rep2
2i mESCs rep1
2i mESCs rep2
Eed-KO mESCs rep1
Eed-KO mESCs rep2
Normalized Capture Hi-C tags (100 DpnII)
H
ELRI contacts
Serum_rep1
Serum_rep2
Eed-KO_rep2
2i_rep2
2i_rep1
Eed-KO_rep1
Day3_rep1
Day1_rep1
Day1_rep2
Day3_rep2
I
Interactions between H3K27me3 marked loci (excluding ELRI contacts)
R=0.840
2i log2 (tags)
Serum log2 (tags)

Figure S2

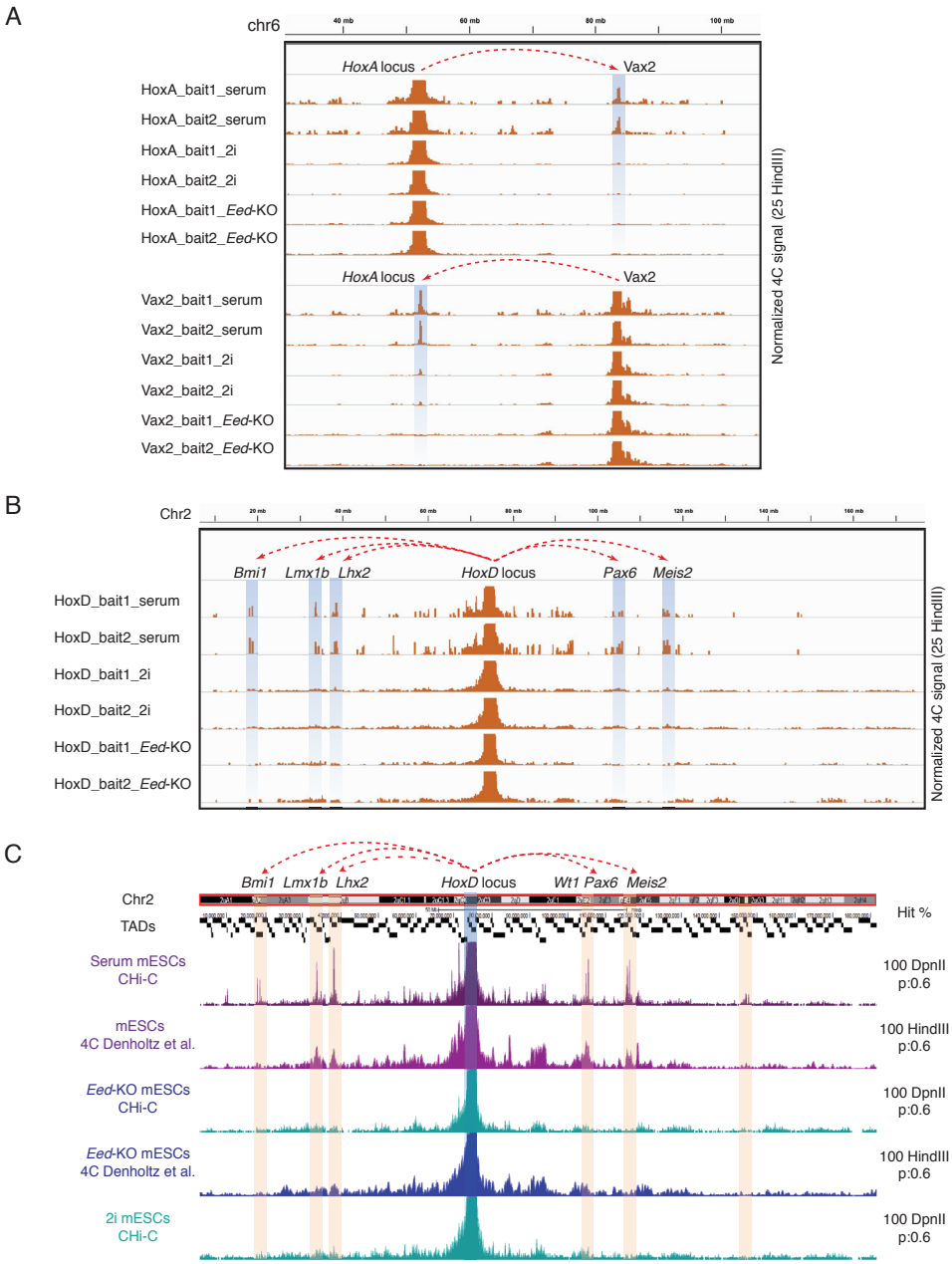

A
chr6
HoxA locus
Vax2
HoxA_bait1_serum
HoxA_bait2_serum
HoxA_bait1_2i
HoxA_bait2_2i
HoxA_bait1_Eed-KO
HoxA_bait2_Eed-KO
Vax2_bait1_serum
Vax2_bait2_serum
Vax2_bait1_2i
Vax2_bait2_2i
Vax2_bait1_Eed-KO
Vax2_bait2_Eed-KO
Normalized 4C signal (25 HindIII)
B
Chr2
Bmi1 Lmx1b Lhx2
HoxD locus
Pax6 Meis2
HoxD_bait1_serum
HoxD_bait2_serum
HoxD_bait1_2i
HoxD_bait2_2i
HoxD_bait1_Eed-KO
HoxD_bait2_Eed-KO
Normalized 4C signal (25 HindIII)
C
Bmi1 Lmx1b Lhx2
HoxD locus
Wt1 Pax6 Meis2
Chr2
TADs
Hit %
Serum mESCs CHi-C
100 DpnII p:0.6
mESCs 4C Denholtz et al.
100 HindIII p:0.6
Eed-KO mESCs CHi-C
100 DpnII p:0.6
Eed-KO mESCs 4C Denholtz et al.
100 HindIII p:0.6
2i mESCs CHi-C
100 DpnII p:0.6

Figure S4

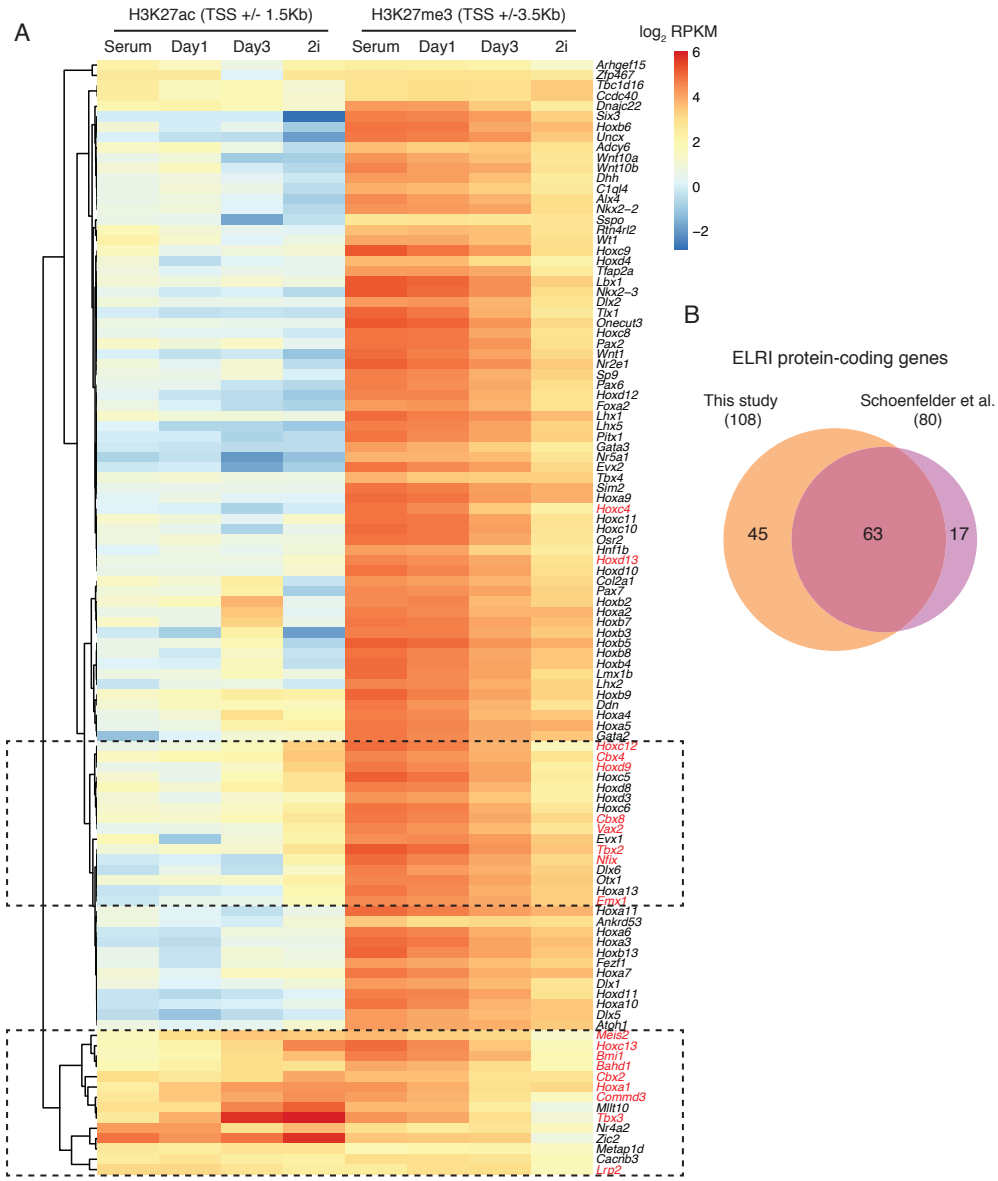
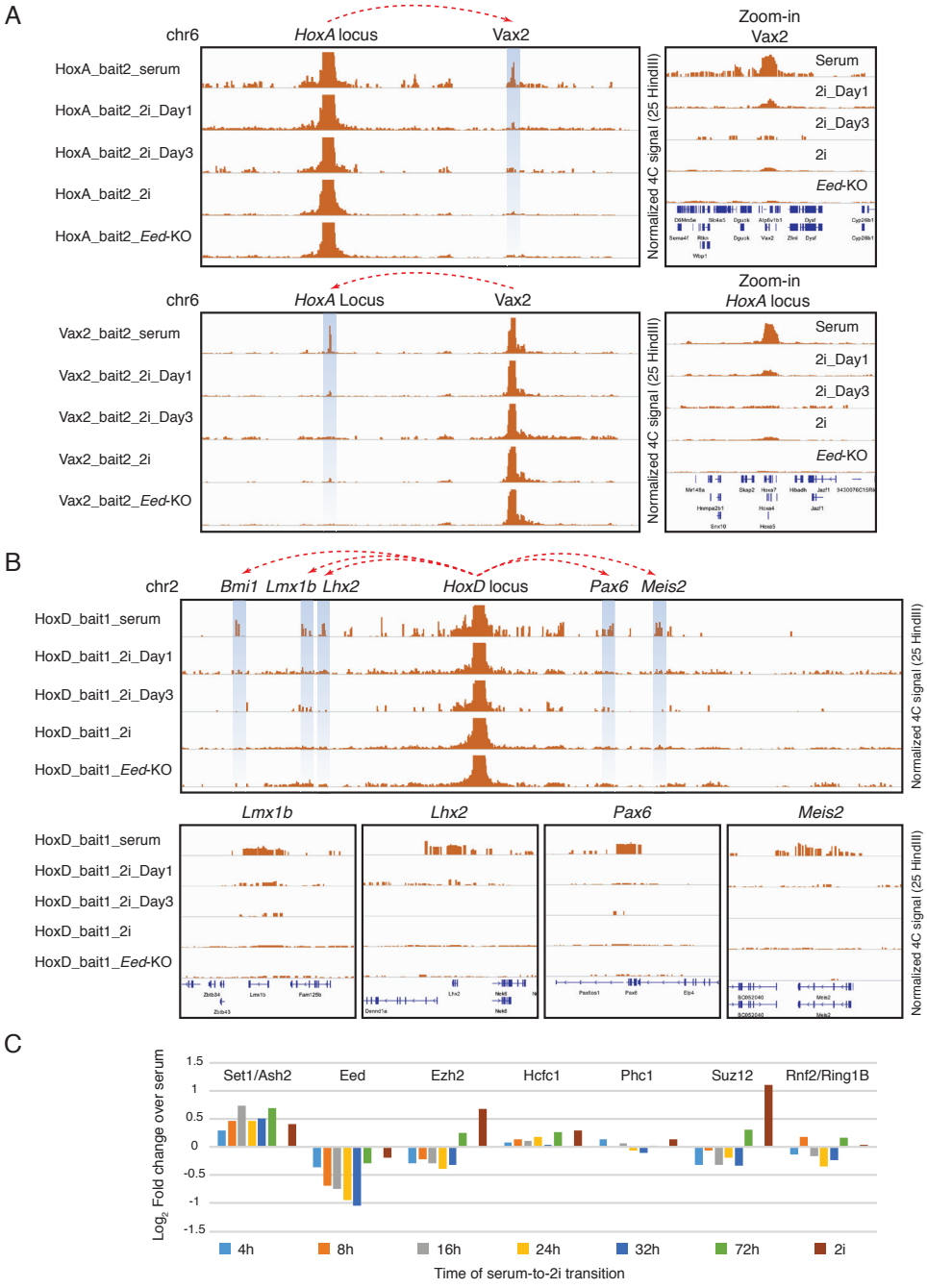

A
H3K27ac (TSS +/- 1.5Kb)
H3K27me3 (TSS +/-3.5Kb)
Serum
Day1
Day3
2i
log2 RPKM
6
4
2
0
−2
Arhgef15
Zfp467
Tbc1d16
Ccdc40
Dnajc22
Six3
Hoxb6
Uncx
Adcy6
Wnt10a
Wnt10b
Dhh
C1ql4
Alx4
Nkx2−2
Sspo
Rtn4rl2
Wt1
Hoxc9
Hoxd4
Tlap2a
Lbx1
Nkx2−3
Dlx2
Tlx1
Onecut3
Hoxc8
Pax2
Wnt1
Nr2e1
Sp9
Pax6
Hoxd12
Foxa2
Lhx1
Lhx5
Pitx1
Gata3
Nr5a1
Evx2
Tbx4
Sim2
Hoxa9
Hoxc4
Hoxc11
Hoxc10
Osr2
Hnf1b
Hoxd13
Hoxd10
Col2a1
Pax7
Hoxb2
Hoxa2
Hoxb7
Hoxb3
Hoxb5
Hoxb8
Hoxb4
Lmx1b
Lhx2
Hoxb9
Ddn
Hoxa4
Hoxa5
Gata2
Hoxc12
Cbx4
Hoxd9
Hoxc5
Hoxd8
Hoxd3
Hoxc6
Cbx8
Vax2
Evx1
Tbx2
Nfix
Dlx6
Otx1
Hoxa13
Epx1
Hoxa11
Ankrd53
Hoxa6
Hoxa3
Hoxb13
Fezf1
Hoxa7
Dlx1
Hoxd11
Hoxa10
Dlx5
Atoh1
Meis2
Hoxc13
Bmi1
Bahd1
Cbx2
Hoxa1
Commd3
Mllt10
Tbx3
Nr4a2
Zic2
Metap1d
Cacnb3
Lrp2
B
ELRI protein-coding genes
This study
(108)
Schoenfelder et al.
(80)
45
63
17

Figure S3


A
chr6
HoxA locus
Vax2
HoxA_bait2_serum
HoxA_bait2_2i_Day1
HoxA_bait2_2i_Day3
HoxA_bait2_2i
HoxA_bait2_Eed-KO
Zoom-in
Vax2
Serum
2i_Day1
2i_Day3
2i
Eed-KO
Normalized 4C signal (25 HindIII)
chr6
HoxA Locus
Vax2
Vax2_bait2_serum
Vax2_bait2_2i_Day1
Vax2_bait2_2i_Day3
Vax2_bait2_2i
Vax2_bait2_Eed-KO
Zoom-in
HoxA locus
B
chr2
Bmi1 Lmx1b Lhx2
HoxD locus
Pax6 Meis2
HoxD_bait1_serum
HoxD_bait1_2i_Day1
HoxD_bait1_2i_Day3
HoxD_bait1_2i
HoxD_bait1_Eed-KO
Normalized 4C signal (25 HindIII)
Lmx1b
Lhx2
Pax6
Meis2
C
Log2 Fold change over serum
1.5
1
0.5
0
-0.5
-1
-1.5
Set1/Ash2
Eed
Ezh2
Hcfc1
Phc1
Suz12
Rnf2/Ring1B
4h
8h
16h
24h
32h
72h
2i
Time of serum-to-2i transition

SUPPLEMENTAL FIGURE LEGENDS

Figure S1, related to Figure 1**.** Characteristics of the DHSs for probe design and reproducibility between CHi-C biological duplicates. (**A**) Venn diagram for DHSs in serum and 2i mESCs. CHi-C probes were designed against the union. (**B**) Genomic distribution of DHSs based on gene annotation and ChIP-seq of histone marks (**C**) SmoothScatter plots show the numbers of di-tags captured by each probe between experiments with high reproducibility. (**D**) SmoothScatter plots show the number of tags for each fragment pairs (fragments of 20kb binned genome) between biological replicates. (**E**) SmoothScatter plots show the number of tags for each called interaction between biological replicates, signifying high reproducibility. (**F**) Both top and bottom figures show high robustness between the normalized tag approach and Hit% approach. Top and bottom figures were generated using a sliding window of 30 DpnII and 100 DpnII fragments, respectively. Top and bottom figures together show high robustness when using different sizes of sliding window. (**G**) Virtual-4C plots showing ELRIs observed on chromosome 2 with high reproducibility between biological duplicates. The rate of the loss of ELRIs in the second experiment (rep2) for serum-to-2i transition was observed to be higher. (**H**) Heat-map of correlations for ELRI contacts between all CHi-C samples. (**I**) Non-ELRI contacts between H3K27me3 marked regions largely unchanged in serum and 2i mESCs.

Figure S2, related to Figure 2**.** Comparison between CHi-C, in-lab generated 4C and 4C from public data sets. (**A**) 4C validation of differential ELRI contact from *HoxA* locus to *Vax2* and reverse, using 2 different baits for serum, 2i and *Eed*$^{-/-}$ mESCs. (**B**) 4C validation of differential ELRI contacts from *HoxA* locus to other ELRI genes highlighted using blue shaded area. (**C**) High correlation visualized between ELRIs detected by CHi-C and publically available 4C data (Denholtz et al., 2013) for serum mESCs and *Eed*$^{-/-}$ mESCs. ELRIs detected from *HoxD* to other loci on chromosome 2 are highlighted by red bands. All ELRI contacts are represented as red arches.

Figure S3, related to Figure 3**.** In-lab generated 4C validation for ELRIs detected by CHi-C on the complete serum-to-2i transition and *Eed*$^{-/-}$ mESCs. (**A**) Loss of the interaction between *HoxA* and *Vax2* revealed by CHi-C validated using 4C experiments on serum-to-2i transition and *Eed*$^{-/-}$ mESCs. (**B**) Loss of the interaction between *HoxD* and other ELRI loci revealed by CHi-C validated using 4C experiments on serum-to-2i transition and *Eed*$^{-/-}$ mESCs. All ELRI contacts are represented as red arches. (**C**) Bar plot showing the changes of whole cell protein levels of PRC components detected using quantitative Mass Spectroscopy for serum-to-2i transition. Y-axis represents Log2 fold change compared to serum mESCs. X-axis represents the time points during serum-to-2i transition (4hrs, 8hrs, 16hrs, 24hrs, 32hrs and 18Days/2i mESCs after medium switch).

Figure S4, related to Figure 4**.** Switching of chromatin state of ELRI promoters from serum to 2i pluripotent state. (**A**) Heat-map shows a gradual loss of H3K27me3 mark on Day1 and Day3 of serum-to-2i transition on all ELRI promoters. A subset of genes (black dash-lined boxes) show a gradual gain in H3K27ac during the transition. All ELRI genes with increased transcriptional activity in 2i form a part of the subset and are labeled in red. (**B**) Venn diagram showing high overlap between ELRI loci detected by our DHS-based CHi-C and promoter-based CHi-C in a parallel study (Schoenfelder et al., 2015).

SUPPLEMENTAL TABLE LEGENDS

Table S1, related to Figure 1. Experimental information of Capture Hi-C and 4C. The union of DHSs in serum and 2i mESCs; all DpnII fragments targeted by CHi-C Probes; list of all virtual baits used; sequencing statistics of CHi-C; list of all primers used for 4C.

Table S2, related to Figure 1. Identification of ELRIs and annotation of ELRI genes. List of ELRI regions from CHiCAGO pipeline and virtual 4C analysis; list of all ELRI genes; all homeobox containing ELRI genes; all ELRI transcription factors;

Table S3, related to Figure 4. ELRI genes that up-regulated in 2i. List of all ELRI genes with significantly increased transcriptional activity in 2i.

EXTENDED EXPERIMENTAL PROCEDURES

DNase I-seq

DNase I libraries were prepared as described in (Shen et al., 2002). Nuclei were isolated using Buffer A (15 mM NaCl, 60 mM KCl, 1 mM EDTA, pH 8.0, 0.5 mM EGTA, pH 8.0, 15 mM Tris-HCl, pH 8.0, 0.5 mM Spermidine) supplemented with 0.015 % IGEPAL CA-630 detergent. DNase I treatment was done for 3 minutes and the reaction was stopped with stop buffer (50 mM Tris-HCl, pH 8, 100 mM NaCl, 0.10 % SDS, 100 mM EDTA, pH 8.0, 1 mM Spermidine, 0.3 mM Spermine). The sample was further fractionated on 9% Sucrose gradient for 24 hours at 25000 rpm at 16 °C. Fractions containing fragments smaller than 1kb were purified and processed according to the Illumina library preparation protocol.

ss-DNA probe design for CHi-C

ss-DNA probes were designed by Roche NimbleGen Inc. as a custom design. Probes were generated for the union of DNase I hotspots (~100,000) for serum and 2i mESCs (~250,000 tiled individual probes), hence providing coverage for all of the open chromatin (Table S1). The length of each probe was on an average 75 nucleotides. The probes were biotinylated for easy capture with streptavidin beads. Total size of the capture probe library was 41.9Mb.

ChIP-seq

ChIP experiments were carried out according to the protocol described (Marks et al., 2012). For Ring-1B and Suz12 ChIPs, 6 million cells per ChIP were used; for H3K27me3, H3K4me1 and H3K27ac ChIPs, 1 million cells per ChIP were used. The antibodies used are as follows: Anti-Ring1B (D22F2) antibody from Cell Signaling and 5µl was used per ChIP; Anti-Suz12 (ab1207) antibody from Abcam and 3µl was used per ChIP; Anti-H3K27me3 (07-449) antibody from Upstate (Millipore) and 3µl was used per ChIP; Anti-H3K27ac (C15410196-10) and Anti-H3K4me1 (CS-037-100) antibodies from Diagenode and 3µl was used per ChIP.

RNA-seq

Total RNA was extracted using RNeasy mini kit (QIAGEN) following the manufacturer's instructions. 5ug of extracted RNA was depleted from ribosomal RNA using Ribo-Zero Gold Kit (Epicentre Madison, Winsconsin, USA). Then, rRNA-depleted RNA was used for library preparation using the TruSeq RNA Sample Prep Kit (Illumina) following the manufacturer's instructions. Libraries were indexed using NEXTflex adapters (Bioo-Scientific Corporation, Austin, TX, USA) and 43bp paired-end sequencing was performed on Illumina instruments using TruSeq reagents (Illumina, San Diego, CA, USA), according to manufacturer's instructions.

4C-Seq

4C assays were performed as described previously (Splinter et al., 2012) with minor modifications. Briefly, 10 million cells were cross-linked for 10 minutes with 2% paraformaldehyde, quenched with glycine and lysed in 50 ml

lysis buffer (10mM Tris pH 7.5, 10mM NaCl, 2% NP-40, 1X protease inhibitors) for 30 minutes. Nuclei were then digested with 800U HindIII enzyme followed by 4 hours 'in nuclei' ligation at 16˚ C with 2000U T4 ligase (NEB) (Nagano et al., 2015). Reverse crosslinked and purified DNA was further digested with 50U DpnII enzyme, followed by circularization. 3200ng of 4C library was amplified with bait-specific inverse primers (Table S1), pooled and purified. Amplified library was adaptor ligated, PCR amplified (8 cycles) and paired-end sequenced on the Illumina NextSeq 500 to obtain 50bp x2 long reads.

Quantitative Mass Spectrometry

Total proteins for each sample were isolated and tryptic digested following a published label-free proteomics protocol (Liu et al., 2012). Three measurements were made for each sample. At least two biological replicates were carried out for each time point. Protein identification and quantification were performed using MaxQuant software (version 1.3.5.7) with standard settings (Cox and Mann, 2008) and searched against the UniProtKB/Swiss-Prot mouse database (generated from version 06-2012).

DNase I-seq and ChIP-seq data analysis

Reads were mapped to the reference mouse genome (mm9) using BWA (Li, 2014) with default parameters and only uniquely mapped reads were kept. PCR duplicates were removed. MACS2 (Zhang et al., 2008) was used to call peaks with parameters "--nomodel --broad". Peaks were filtered based on peak score. Overlap peaks were merged if they are from ChIP-seq samples of the same antibody or DNase I-seq samples.

RNA-seq data analysis

MMSEQ package (Turro et al., 2011) was used to infer gene expression levels. Reads were mapped to mouse gene annotation (Ensemble release 67). MMDIFF (Turro et al., 2014) was used to calculate probability of differential expression with default settings. Genes were considered to be up regulated in 2i, if RPKM > 1 in each 2i biological replicate and probability > 0.2.

4C-Seq Data analysis

To improve the mappability of the sequencing reads, we generated a reduced genome by extracting the sequences flanking the HindIII cutting sites (30bp on each strand from the HindIII cutting sites to downstream) based on the reference mouse genome (mm9). Then we evaluated the mappability of the extracted sequences (each strand separately) and only uniquely mappable HindIII cutting sites were considered for downstream analysis.

All the reads from each library were parsed based on the bait-specific primer sequence and mapped to the reduced genome using BWA with the default parameters. 4C signals and 4C hits were calculated using a sliding window of a fixed number of HindIII fragments and normalized to the total number of uniquely mapped reads.

compared to when HindIII is used as the restriction enzyme. To achieve a 3D map at single DpnII resolution, innumerous reads will be needed. Therefore, to lower the need of sequencing depth, we created "virtual targeted fragments" by merging adjacent targeted DpnII fragments and "virtual non-targeted fragments" using tiles of 5 DpnII fragments. Besides the default threshold of 5 for the CHiCAGO score, we also called for at least 5 reads to support an interaction.

Identification of ELRIs

We defined *Hox* related extremely long-range interactions as ELRIs. Hence, we used the loci of the four *Hox* gene clusters as the viewpoints separately and did comprehensive identification of ELRI genes with a 4C-like analysis approach (Van De Werken et al., 2012). Firstly, we extracted the reads that one-end is from a given point of view (one of the four *Hox* gene clusters). Secondly, CHi-C signal was calculated using the numbers of di-tags in a sliding window of a 20 DpnII fragments and normalized to sequencing depth; we also calculated the percentage of hits in the sliding window. ELRIs are extremely long-range interactions, hence the intrinsic background from CHi-C is virtually non-existent at these interacting loci. Considering the noise, a stringent cutoff for both CHi-C signal (at least 5) and percentage of hits (at least 15%) were used to identify the candidates of ELRI loci. We filtered for differential long-range interactions between serum and 2i mESCs by calling for the sum of normalized di-tags in a sliding window of 20 DpnII fragments with at least three fold changes. All the above parameters were used for the identification of ELRIs.

Calculation of the strength of chromatin contact

We defined the chromatin contact strength as the number of di-tags for the contact between two genomic fragments normalized to sequencing depth (100M informative reads), the length of the first fragment and the length of the second fragment. To avoid zero, 0.01 was added to the strength before performing $\log_2$.

SUPPLEMENTAL REFRENCES

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Genovese, C.R., Roeder, K., and Wasserman, L. (2006). False discovery control with p-value weighting. Biometrika *93*, 509–524.

Li, H. (2014). Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 1–9.

Liu, N.Q., Braakman, R.B.H., Stingl, C., Luider, T.M., Martens, J.W.M., Foekens, J. a., and Umar, A. (2012). Proteomics pipeline for biomarker discovery of laser capture microdissected breast cancer tissue. J. Mammary Gland Biol. Neoplasia *17*, 155–164.

Nagano, T., Várnai, C., Schoenfelder, S., Javierre, B.-M., Wingett, S.W., and Fraser, P. (2015). Comparison of Hi-C results using in-solution versus in-nucleus ligation. Genome Biol. *16*, 175.

Shen, M., Kawamoto, T., Teramoto, M., Makihira, S., Fujimoto, K., Yan, W., Noshiro, M., and Kato, Y. (2001). Induction of basic helix-loop-helix protein DEC1 (BHLHB2)/Stra13/Sharp2 in response to the cyclic adenosine monophosphate pathway. Eur. J. Cell Biol. *80*, 329–334.

Shen, M., Yoshida, E., Yan, W., Kawamoto, T., Suardita, K., Koyano, Y., Fujimoto, K., Noshiro, M., and Kato, Y. (2002). Basic helix-loop-helix protein DEC1 promotes chondrocyte differentiation at the early and terminal stages. J. Biol. Chem. *277*, 50112–50120.

Splinter, E., de Wit, E., van de Werken, H.J.G., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. Methods *58*, 221–230.

Turro, E., Su, S.-Y., Gonçalves, Â., Coin, L.J.M., Richardson, S., and Lewin, A. (2011). Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. Genome Biol. *12*, R13.

Turro, E., Astle, W.J., and Tavaré, S. (2014). Flexible analysis of RNA-seq data using mixed effects models. Bioinformatics *30*, 180–188.

Van De Werken, H.J.G., De Vree, P.J.P., Splinter, E., Holwerda, S.J.B., Klous, P., De Wit, E., and De Laat, W. (2012). 4C technology: Protocols and data analysis. Methods Enzymol. *513*, 89–112.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Chapter 4

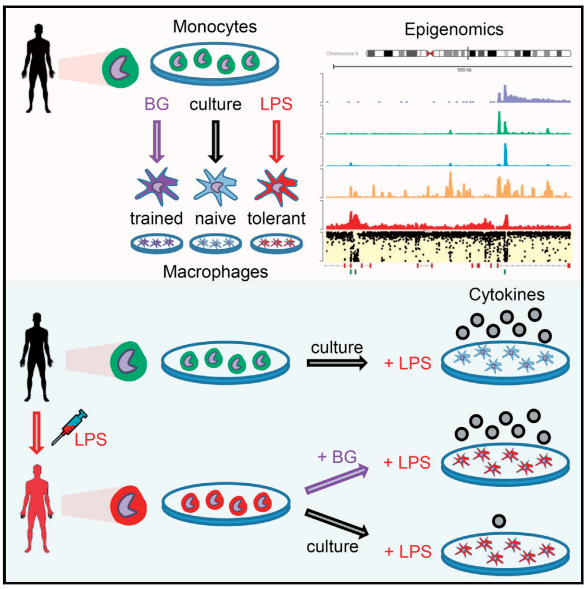
β-Glucan Reverses the Epigenetic State of LPS- Induced Immunological Tolerance

Article

Cell

β-Glucan Reverses the Epigenetic State of LPS-Induced Immunological Tolerance

Graphical Abstract



Authors

Boris Novakovic, Ehsan Habibi, Shuang-Yin Wang, ..., Joost H.A. Martens, Colin Logie, Hendrik G. Stunnenberg

Correspondence

h.stunnenberg@ncmls.ru.nl

In Brief

As part of the International Human Epigenome Consortium (IHEC), this study reveals that β-glucan reverses the state of epigenetic immune tolerance that develops after exposure to LPS and restores the ability of human macrophages to produce cytokines that are critical for anti-pathogen responses. Explore the Cell Press IHEC webportal at http://www.cell.com/consortium/IHEC.

Highlights

- Epigenetic and transcriptional characterization of human macrophage tolerance
- LPS-exposed monocytes fail to induce macrophage-specific downstream pathways
- Monocyte-induced tolerance of macrophages can be reversed by β-glucan at the epigenetic level
- In-vivo-tolerized monocytes can be reverted to a responsive phenotype ex vivo by β-glucan

Data Resources

GSE85246
GSE85243
GSE85245
GSE87218
EGAD00001002693

Novakovic et al., 2016, Cell *167*, 1354–1368
November 17, 2016
http://dx.doi.org/10.1016/j.cell.2016.09.034

Cell

Article

β-Glucan Reverses the Epigenetic State of LPS-Induced Immunological Tolerance

Boris Novakovic,[1,7] Ehsan Habibi,[1,7] Shuang-Yin Wang,[1,7] Rob J.W. Arts,[2] Robab Davar,[1] Wout Megchelenbrink,[1] Bowon Kim,[1] Tatyana Kuznetsova,[1] Matthijs Kox,[3] Jelle Zwaag,[3] Filomena Matarese,[1] Simon J. van Heeringen,[4] Eva M. Janssen-Megens,[1] Nilofar Sharifi,[1] Cheng Wang,[1] Farid Keramati,[1] Vivien Schoonenberg,[1] Paul Flicek,[5] Laura Clarke,[5] Peter Pickkers,[3] Simon Heath,[6] Ivo Gut,[6] Mihai G. Netea,[2] Joost H.A. Martens,[1] Colin Logie,[1] and Hendrik G. Stunnenberg[1,8,*]

[1]Department of Molecular Biology, Faculty of Science, Radboud University, 6525 GA Nijmegen, the Netherlands
[2]Department of Internal Medicine, Radboud University Medical Center, Radboud Center for Infectious Diseases (RCI), 6525 GA Nijmegen, the Netherlands
[3]Department of Intensive Care Medicine, Radboud University Medical Center, Radboud Center for Infectious Diseases (RCI), 6500 HB Nijmegen, the Netherlands
[4]Department of Molecular Developmental Biology, Faculty of Science, Radboud University, 6525 GA Nijmegen, the Netherlands
[5]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK
[6]Centro Nacional de Análisis Genómico (CNAG), Parc Científic de Barcelona, 08028 Barcelona, Spain
[7]Co-first author
[8]Lead Contact
*Correspondence: h.stunnenberg@ncmls.ru.nl
http://dx.doi.org/10.1016/j.cell.2016.09.034

SUMMARY

Innate immune memory is the phenomenon whereby innate immune cells such as monocytes or macrophages undergo functional reprogramming after exposure to microbial components such as lipopolysaccharide (LPS). We apply an integrated epigenomic approach to characterize the molecular events involved in LPS-induced tolerance in a time-dependent manner. Mechanistically, LPS-treated monocytes fail to accumulate active histone marks at promoter and enhancers of genes in the lipid metabolism and phagocytic pathways. Transcriptional inactivity in response to a second LPS exposure in tolerized macrophages is accompanied by failure to deposit active histone marks at promoters of tolerized genes. In contrast, β-glucan partially reverses the LPS-induced tolerance in vitro. Importantly, ex vivo β-glucan treatment of monocytes from volunteers with experimental endotoxemia re-instates their capacity for cytokine production. Tolerance is reversed at the level of distal element histone modification and transcriptional reactivation of otherwise unresponsive genes.

INTRODUCTION

Accumulating evidence suggests that monocytes can be reprogrammed by exposure to microbe-associated molecular patterns (MAMPs) during their time in the circulation (Quintin et al., 2014). In this model, immune tolerance in myeloid cells, be they monocytes in the circulation or macrophages in the tissues (lipopolysaccharide macrophages [LPS-Mfs]), represents one extreme in the spectrum of innate immune memory and can be induced by high bacterial burden in vivo or lipopolysaccharide (LPS) exposure in vitro (Netea et al., 2016). On the other hand, trained immunity can be induced by exposure to certain vaccines, microbial components, or metabolites, and is a state characterized by increased pro-inflammatory response to secondary unrelated infections (Netea et al., 2016). We recently showed that tolerance (induced by LPS) and trained immunity (induced by *Candida albicans* β-glucan [BG]) are both associated with specific epigenomic states (Cheng et al., 2014; Saeed et al., 2014). Most notably, the identity of these macrophage subtypes was specified by differences in primed and active distal element repertoires (Saeed et al., 2014).

Monocytes and macrophages play an important role in the pathophysiology of sepsis and inflammation, along with other innate and adaptive immune cells (Biswas and Lopez-Collazo, 2009). Transcriptome analysis of tolerant monocytes from sepsis patients (Shalova et al., 2015) and a mouse sepsis model (Foster et al., 2007) reveals that the tolerized phenotype cannot be explained purely through failure of specific signaling pathways induced by pattern recognition receptors to activate downstream genes. This implicates a role for local chromatin architecture and specific transcriptional regulators in controlling the expression of tolerized genes (Glass and Natoli, 2016). Further, studies in human cancers have revealed commonalities between inflammation and cancer associated tolerance, including the role for IDO1 in both (Bessede et al., 2014). Accordingly, several anti-cancer drugs, such as bromodomain and extraterminal domain family (BET) inhibitors and a topoisomerase inhibitor, have proven efficacious in blocking inflammation-associated death in mice (Nicodeme et al., 2010; Rialdi et al., 2016). The specific epigenetic and transcriptional remodeling induced by the initial LPS exposure and the extent to which it specifies tolerance to future LPS exposure are unknown.

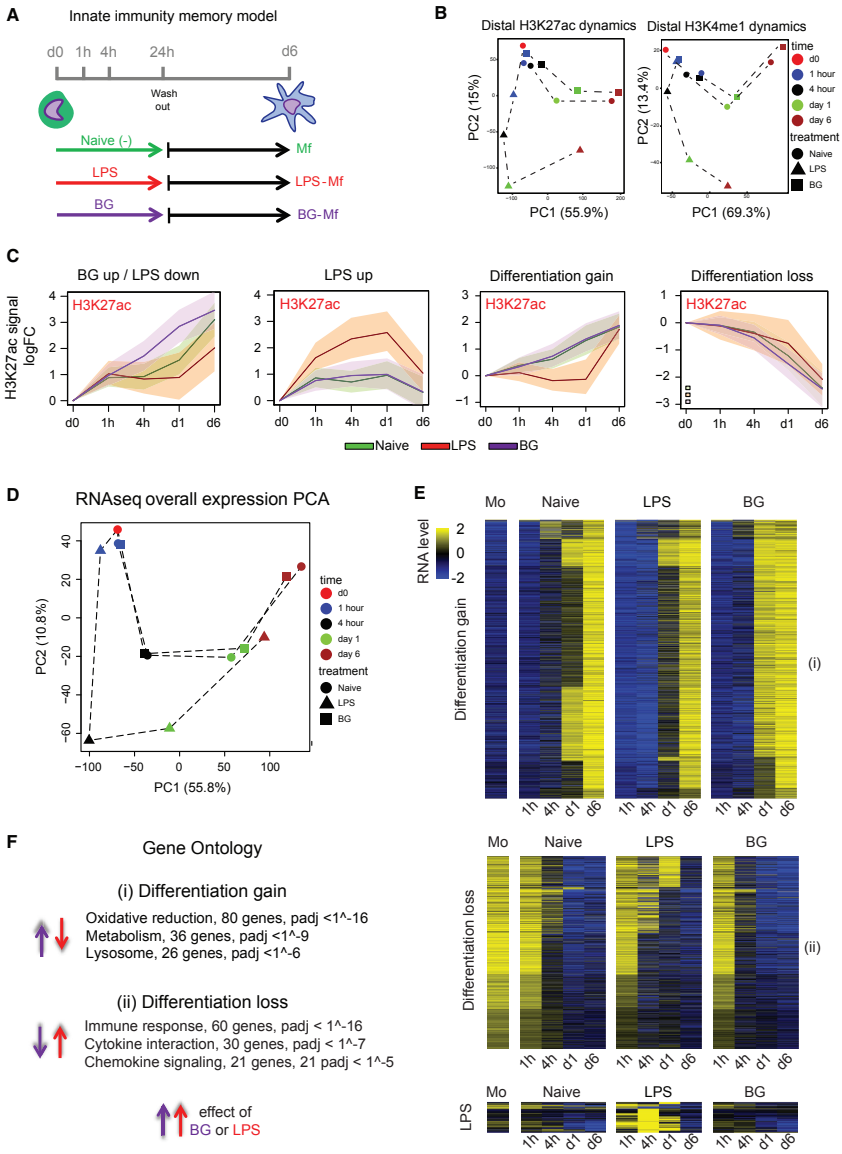


Figure 1. Epigenomic and Transcriptomic Remodeling of Monocytes Induced by Exposure to LPS or BG

(A) Experimental setup for epigenomic interrogation of monocyte-to-macrophage differentiation and induction of tolerance (with LPS) or trained immunity (with BG). (B) PCA plots of H3K27ac and H3K4me1 dynamic enhancers (monocytes, red circle; naive, circle; LPS, triangle; BG, square; 1 hr, blue; 4 hr, black; day 1, green; and day 6, brown). Dynamic H3K27ac patterns show a clear deviation from the differentiation pathway (PC1) in LPS-treated cells. On the other hand, BG-treated cells at day 1 are well on their way toward a full macrophage epigenetic profile.

(legend continued on next page)

Here, as part of the BLUEPRINT epigenome consortium (http://www.blueprint-epigenome.eu), we report the time-resolved, comprehensive epigenomes of human monocyte-to-macrophage differentiation and induction of tolerance with LPS and training with BG. Our epigenomic analysis revealed that tolerance and trained immunity involve opposing regulation of common pathways during early exposure to MAMPs, leading to distinct epigenomic states in the two macrophage subtypes. We therefore hypothesized that BG may be capable of reversing LPS-induced tolerance. We show that ex vivo BG exposure can reinstate a responsive phenotype in both monocytes tolerized by ex vivo LPS exposure and monocytes tolerized by in vivo experimental endotoxemia in healthy volunteers. This reversal of tolerance involves epigenomic reprogramming of macrophages.

RESULTS

Distinct Temporal Epigenetic Remodeling in Response to Microbial Components

Two innate immune memory states can be induced in culture through an initial exposure of primary human monocytes to either LPS or BG for 24 hr, followed by removal of stimulus and differentiation to macrophages for an additional 5 days (Figures 1A and S1; Quintin et al., 2012; Saeed et al., 2014). The three subtypes of macrophages generated in this study were untreated naive macrophages (naive-Mfs), LPS-exposed tolerized macrophages (LPS-Mfs), and BG-exposed trained macrophages (BG-Mfs). To gain insight into the mechanisms and order of events that ultimately lead to these three subtypes, we generated epigenomic data at several time points during this process (two donors; summarized in Table S1 and Figure S1; GEO: GSE85246). Depending on the modification, 2%–31% of marked regions showed dynamics during differentiation or LPS or BG exposure, with H3K27ac at promoters and enhancers being the most dynamic mark in number and range (Figure S1; Table S1). Interestingly, epigenetic changes were observable as early as 1 hr in response to LPS and 4 hr to BG (Figure 1B). The overall H3K27ac pattern at dynamic promoters and enhancers indicates that the most pronounced changes are associated with differentiation (principal component 1 [PC1]), with BG- and RPMI-treated monocytes partially establishing macrophage-specific active regions already by day 1 (Figure 1B). Conversely, LPS treatment results in establishment of pro-inflammatory associated active elements (PC2) and stunted differentiation, followed by partial "catch up" establishment of differentiation marks following removal of stimulus (Figures 1B and S1). Contrary to this catch up of H3K27ac marked enhancers, the H3K4me1 marked enhancer repertoire of LPS-Mfs is significantly different from those of naive-Mfs and BG-Mfs (Figure 1B). Repressive marks, H3K9me3 and H3K27me3, showed no dynamics during the first 24 hr, indicating little role in the early, priming phase of innate immune memory (Figure S1C; Table S1).

In total, 17,500 enhancers with dynamic H3K27ac were identified (Figure S1B). The two largest clusters show gain (n = 4,028) or loss (n = 6,462) of H3K27ac during differentiation in all three macrophage subtypes (Figure 1C and S1B). The closest genes associated with differentiation gain or loss clusters are associated with leukocyte differentiation, activation, metabolism, and phagocytosis (Table S2). Upon monocyte exposure to LPS, H3K27ac induction precedes a temporally delayed H3K4me1 (Figures 1C and S1B). The closest genes associated with these enhancers are involved in cytokine response and nuclear factor κB (NF-κB) signaling, among other well-known LPS-response pathways (Table S2). The "BG up/LPS down" enhancer cluster shows accelerated H3K27ac deposition in BG-exposed monocytes and little to no H3K27ac accumulation in LPS-exposed monocytes relative to naive-Mfs (Figure 1C). This cluster is composed of >3,200 enhancers and shows concordant increase in H3K4me1 to day 6 (Figure S1B). Chromatin segmentation analysis using EpicSeq (Mammana and Chung, 2015) revealed that these regions gain H3K4me1 at the expense of repressive H3K27me3 markings in naive-Mfs and BG-Mfs (Figure S1D). Conversely, LPS-Mfs maintain a chromatin state more similar to monocytes, primarily low H3K4me1 with the presence of H3K27me3 (Figure S1D). The closest genes to these enhancers are involved in lipid biosynthesis and lysosome and leukocyte differentiation (Table S2), indicating that BG exposure leads to the accumulation of membrane components necessary for phagocytosis and cytokine release, whereas LPS exposure prevents their activation (Figure 1C).

Transcriptome Changes Modulated by LPS and β-Glucan

RNA sequencing (RNA-seq) was performed on the same time points as epigenetic marks (n = 2 donors; Figure 1A). General kinetics similar to those unveiled for epigenetic remodeling was observable, with monocytes clustering after a short exposure to BG and LPS (Figure 1D). Over the time course, the major changes in gene expression patterns were associated with differentiation (PC1, 55.8% of the variance) and LPS exposure (PC2, 10.8% of the variance), which is most pronounced at 4 hr and day 1 (Figure 1D). Over 5,700 protein-coding genes showed dynamic expression (fold change [FC] > 2, adjusted p value [padj] < 0.05) in our model between either treatments or time points (Table S3; Figures 1E and S2A). LPS-induced genes

(C) A total of 17,500 H3K27ac dynamic gene-distal regions were identified and can be clearly separated into four clusters: BG up/LPS down, LPS up, differentiation gain, and differentiation loss. Solid lines are median log-FC relative to day 0, and shaded areas represent the 25th and 75th quartile. Naive cells are shown as a green line, LPS as a red line, and BG as a purple line. H3K4me1 at these regions can be seen in Figure S1B; LPS induces early H3K27ac accumulation, followed by long-term H3K4me1 marking, while BG induces concurrent accumulation of H3K27ac and H3K4me1.

(D) PCA plots showing the relationships among all samples based on dynamic gene expression. PC1 explains most of the variation and is associated with differentiation. PC2 is LPS related, with LPS 4 hr and LPS day 1 samples separating from the corresponding naive and BG samples.

(E) Heatmap of differentiation associated genes, as well as those induced by LPS or BG exposure. The general trend in expression is that BG exposed cells start to express differentiation associated genes faster (at day 1) than naive cells, while LPS exposed cells lag behind.

(F) Top pathways associated with differentiation and showing opposing directions in response to BG and LPS.

See also Figures S1, S2, and S3 and Tables S1, S2, S3, and S4.

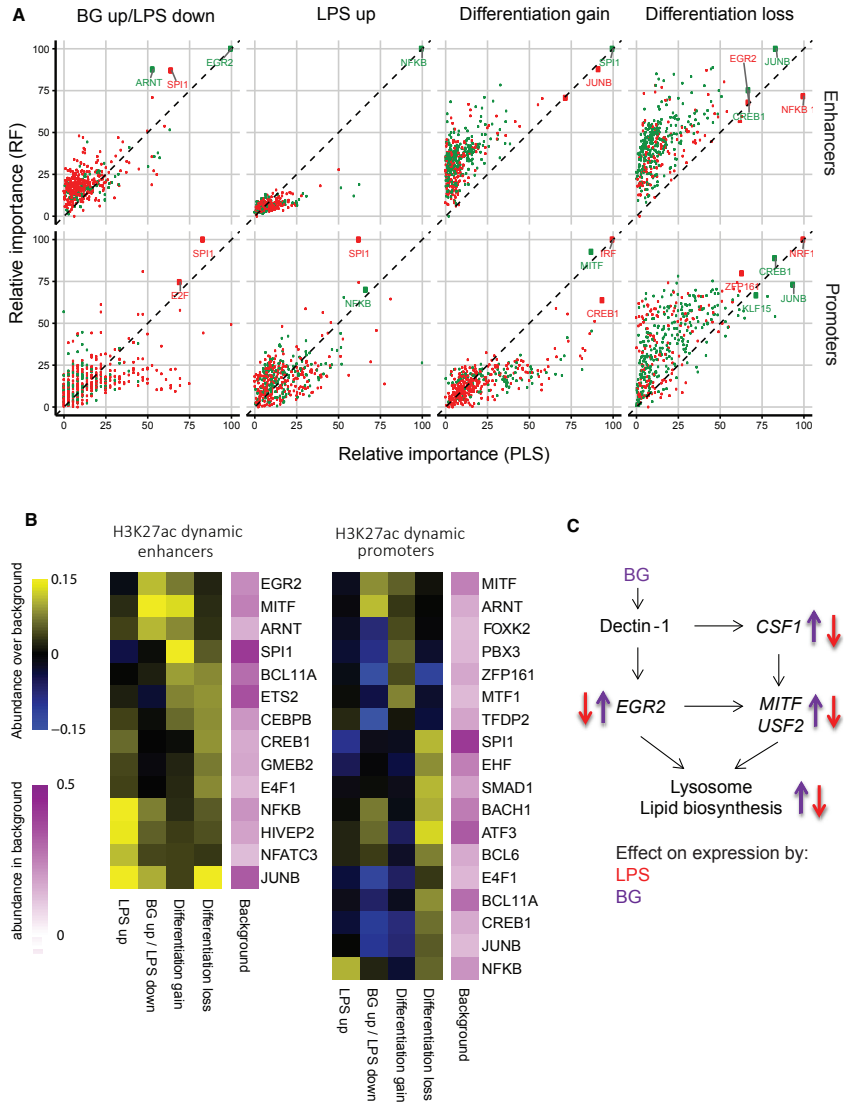


Figure 2. Motif Enrichment at Epigenetically Dynamic Promoters and Enhancers and Associated Transcription Factor Networks

Motif enrichment analysis was performed on ATAC-sequencing (nucleosome-free) peaks that overlap H3K27ac dynamic enhancers and H3K27ac promoters. (A) Random forest (RF) and a partial least-squares (PLS) classifiers were trained using the TF motifs found by GIMME to determine features (TF motifs) based on their ability to separate the 4 H3K27ac clusters shown in Figure 1C. Both classifiers produce a feature importance score (between 0 and 100), which is a measure of how "characteristic" the presence or absence of the TF motif is for the considered cluster. Green dots represent positive features (motif over-represented in cluster), and red dots represent negative features (motif under-represented in cluster). The EGR2 motif was the strongest positive feature for the BG up/LPS down

(legend continued on next page)

were involved in immune response, whereas LPS-delayed genes were generally differentiation associated (Figure 1E; Table S4). The major ontologies of BG-induced genes (662 genes at day 1) were lipid biosynthesis, metabolism, and the lysosome pathway (Figure 1E; Table S4). Intersection between exposure-dependent gene expression and promoter acetylation patterns showed a strong overlap between H3K27ac and gene expression temporal profiles (Figures S2B and S2C).

LPS-Specific DNA (De)methylation Signatures

Recent studies have revealed extensive DNA methylation remodeling during B cell (Kulis et al., 2015) and osteoclast differentiation (de la Rica et al., 2013; Nishikawa et al., 2015). Considering that our ex vivo differentiation model occurs in the absence of cell division, we were interested to see the extent to which (de)methylation plays a role during monocyte-to-macrophage differentiation and innate immune memory. Unlike the comprehensive histone modification remodeling, consistent DNA methylation change (at least 30% change and four or more significant differentially methylated CpGs per differentially methylated region [DMR]) was limited to a few hundred genomic regions (Figure S3). The vast majority of DMRs showed loss of methylation during monocyte-to-macrophage differentiation irrespective of MAMP exposure (Figure S3B), consistent with recent findings in macrophages and dendritic cells (Vento-Tormo et al., 2016). We did not observe a role for DNA methylation in "training" the macrophages for future transcriptional response to infection. More than 90% of DMRs occurred at distal elements marked by H3K4me1, and only 6% occurred at promoters (Figure S3C). Cumulatively, our data indicate that LPS-specific DNA methylation changes occur and, due to the more stable nature of this mark, may represent a useful biomarker for LPS-induced macrophage tolerance (Figures S3D and S3E).

LPS- and β-Glucan-Specific Transcriptional Networks

Motif analysis was used to gain insight into which pathways and transcription factors (TFs) regulate the epigenetic changes associated with differentiation and LPS or BG exposure. Four clusters of enhancers and promoters were designated based on H3K27ac dynamics over time: BG up/LPS down, "LPS up," "differentiation gain," and "differentiation loss" (Figures 1C and S1B). Two classifiers (random forest [RF] and a partial least-squares [PLS]) were trained, using the TF motifs found by GIMME (van Heeringen and Veenstra, 2011). Both score features (TF motifs) were based on their ability to separate the clusters—the so-called feature importance score (between 0 and 100)—which is a measure of how characteristic the presence or absence of the TF motif is for the considered cluster. Both classifiers were trained with the *caret* R-package using 10-fold cross-validation, repeated five times. We define positive (green dots) and negative predictors (red dots) as TF motifs that are more or less abundant of the considered cluster compared to the other clusters, respectively.

Enhancers that show differentiation gain in H3K27ac were enriched for the SPI1 (PU.1) motif, while LPS-induced active enhancers were enriched for NF-κBmotif (Figures 2A and 2B). The top positive predictor motif for the BG up/LPS down cluster was EGR2, with a score of 100, followed by ARNT (Figure 2A). EGR2 is downstream of dectin-1 (Goodridge et al., 2007) and shows prominent, transient induction in BG-exposed monocytes (Figure S4A). Enhancers with EGR2 motifs are mainly associated with genes involved in lipid metabolism and biosynthesis and lysosome function (Figures 1E and 1F). The early activation of these pathways in BG may account for the higher expression of *LAMP1*, the major component of the mature lysosome, in BG-Mfs (Figure S4A).

Interestingly, LPS-exposed monocytes do not transiently activate *EGR2* (Figure S4A). The discordant effect of BG and LPS on *EGR2* expression, the differential H3K27ac deposition at associated enhancers, and expression of downstream lipid metabolism genes suggests that this pathway plays a role in inducing trained immunity as opposed to tolerance. In order to further confirm the relationship between EGR2 and downstream lipid pathways, the DNA-binding motif of EGR2 was scanned at the promoters of known transcription factors, as well as lipid metabolism and lysosome genes that are induced in BG-Mfs compared to monocytes (Figure S4B). The EGR2 motif was found at the promoter of several highly expressed TFs, including MITF, which is a positive identifier for the differentiation gain promoter cluster (Figure 2A), and is also not activated in LPS-exposed monocytes (Figure S4A). Cumulatively, EGR2, MITF, and downstream TF motifs were found at the promoters of 79% of induced lipid metabolism and lysosome genes (Figure 2B). This analysis suggests that BG/dectin-1-induced EGR2 activation leads to higher expression of downstream TFs (e.g., MITF) and the establishment of promoters and enhancers that drive the expression of lysosomal and lipid metabolism genes (Figure 2C). Given the importance of lipid pathways in macrophage function, the opposing effect of LPS and BG on these genes suggests that this pathway may play a critical role in the low cytokine release in LPS-Mfs and elevated release in BG-Mfs.

Transcriptional Response of Tolerized Macrophages to LPS Re-exposure

Previous analysis in an ex vivo mouse model showed that tolerant LPS-Mfs are impaired in their ability to produce pro-inflammatory

enhancer cluster, NF-κB for the LPS up cluster, SPI1 (PU.1) for the differentiation gain cluster, and JUNB for the differentiation loss cluster. At the promoter regions NF-κB was a positive feature for LPS up cluster, MITF for the differentiation gain cluster, and CREB1 and JUNB for the differentiation loss cluster.
(B) Motif enrichment is plotted as absolute difference in abundance compared to background (yellow, higher abundance than background; blue, lower abundance than background) for the top enriched motifs. Consistently identified transcription factor motifs include SPI1 at differentiation associated enhancers, NF-κB at LPS enhancers, and EGR2 and MITF at BG enhancers. Abundance increase over background supports the level of importance score.
(C) A diagram of the transcription factor network based on EGR2 and MITF motif occurrence at BG-induced lysosome and lipid metabolism genes. Purple arrows indicate the direction of expression induced by BG exposure, and red arrows indicate the direction of expression induced by for LPS exposure. BG exposure induces transient expression of the genes, while LPS exposure inhibits activation. The full network based on promoter abundance is shown in Figure S4B.
See also Figure S4.

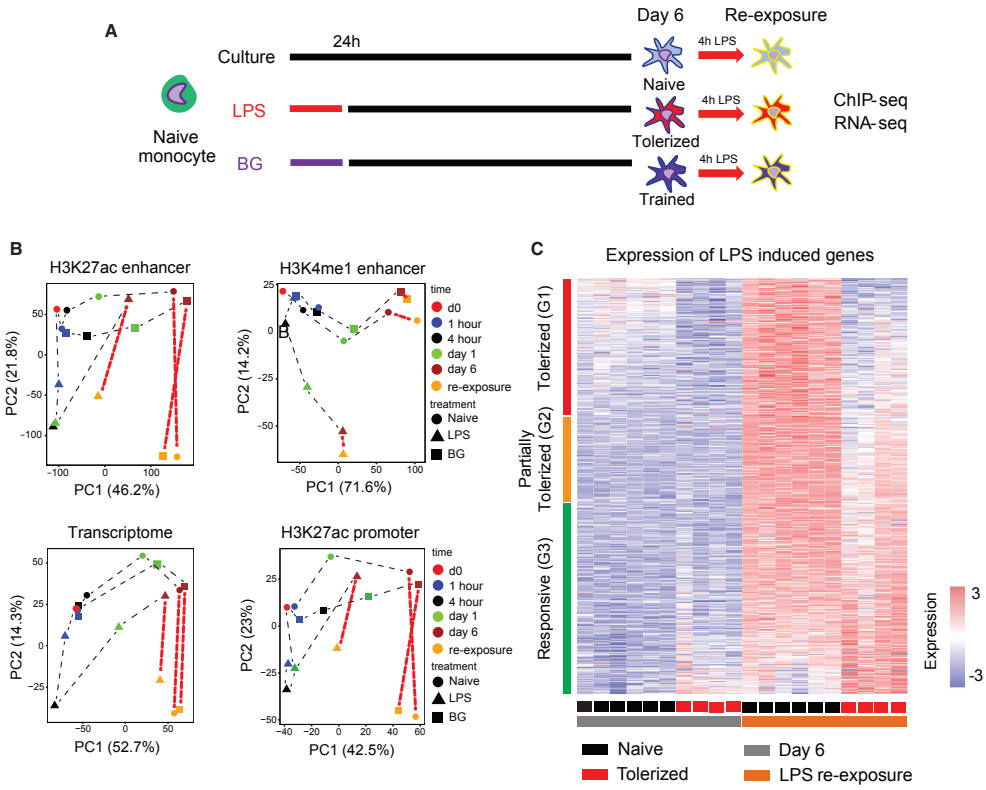


Figure 3. Macrophage Endotoxin Tolerance Defined at the Transcriptional Level following LPS Re-exposure

(A) The innate immune memory model, including data collection at LPS re-exposure at day 6.

(B) PCA plots of dynamic RNA-seq, H3K27ac at promoters and enhancers, and H3K4me1 peaks, including LPS re-exposure samples. After re-exposure to LPS, significant enhancer H3K27ac changes occur in LPS-Mfs, indicating that they are capable of activating their enhancers. However, the level of their response is lower compared to monocytes, naive-Mfs, and BG-Mfs, which can be seen on the second principal component. Unlike RNA and H3K27ac, H3K4me1 does not show significant changes following LPS re-exposure in any of the three macrophage subtypes.

(C) The total macrophage transcriptional response (750 genes) to LPS was separated into three groups based on the induction of genes in LPS-Mfs, relative to naive-Mfs and BG-Mfs, revealing a gradient in LPS-Mf response to LPS re-exposure. The groups are (G1) tolerized genes, (G2) partially tolerized genes, and (G3) responsive genes.

See also Figure S5.

cytokines, but maintain their ability to express other genes, such as those required for tissue repair (Foster et al., 2007). Given the wide-ranging epigenetic alterations in LPS-Mfs (Figure 1C; Table S1), we sought to investigate the epigenetic basis for endotoxin tolerance by exposing differentiated naive-Mfs, LPS-Mfs, and BG-Mfs to LPS for 4 hr (LPS re-exposure) (Figure 3A). The overall transcriptional and histone modification changes induced in macrophages by LPS re-exposure are shown in Figure 3B, and few differences were observed between naive-Mfs and BG-Mfs. LPS-Mfs show an avid response to LPS re-exposure both transcriptionally and with H3K27ac deposition at promoters and distal enhancers (observable as large shift in PC2; Figure 3B). This indicates that tolerized macrophages can and do respond to LPS at the epigenetic and transcriptional level. However, from H3K27ac and H3K4me1 principal-component analysis (PCA), it is clear that the epigenetic profile of LPS-Mfs is markedly different from that of naive-Mfs and BG-Mfs (observable as an LPS-Mfs lag on PC1; Figure 3B).

Polytomous modeling was used to separate genes based on their transcriptional response to LPS re-exposure (4 hr) in macrophages at day 6. In total, 780 genes showed higher expression (FC > 2, posterior probability > 0.3) in naive-Mfs following 4-hr

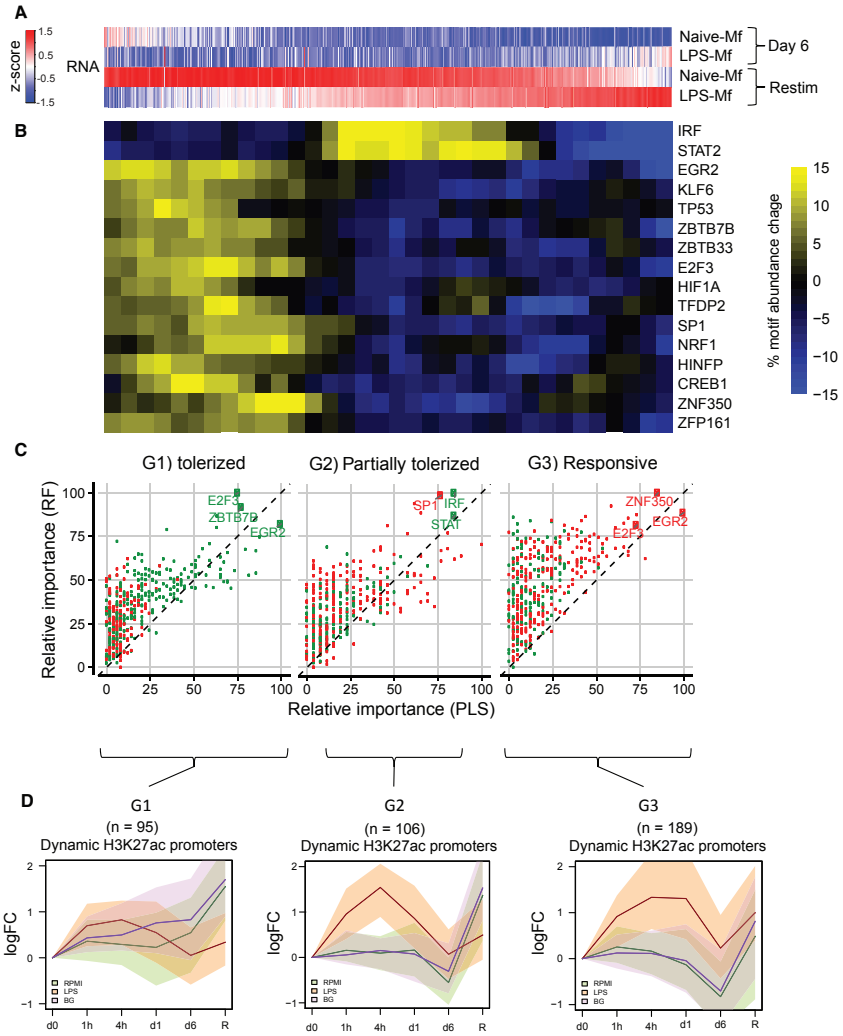


Figure 4. Histone Modification Dynamics and Open Chromatin Analysis at Tolerized Gene Promoters

(A) Heatmap showing average expression of 777 LPS-responsive genes in naive-Mfs. Genes are ranked based on their induction in LPS-Mfs, first by tolerance group (G1, G2, and G3) and then by relative induction compared to naive-Mfs within each group. Response to LPS re-exposure is a gradient in LPS-Mfs, with the most tolerized genes on the left and the most responsive genes on the right.

(B) Heatmap showing abundance of significant motifs in the promoter regions of the three macrophage LPS-responsive gene groups. The tolerized gene promoters are enriched for several transcriptional repressors, such as EGR2 and TP53, while the partially tolerized gene promoters are enriched for IRF and STAT motifs.

(C) Random forest (RF) and a partial least-squares (PLS) classifiers importance score (between 0 and 100) for each tolerized gene cluster (G1, tolerized; G2, partially tolerized; and G3, responsive). Green dots represent over-represented motifs, and red dots under-represented motifs. The top features of G1 gene

(legend continued on next page)

LPS exposure (Figure 3C; Table S5). Transcriptional responsiveness to LPS re-exposure in LPS-Mfs is a gradient, with genes showing complete tolerance (unresponsiveness) (cluster G1), a partial response (G2), or a full response comparable to naive-Mfs (G3) (Figures 3C and S5A). Cytokine genes were the most enriched group and were spread across the LPS re-exposure response gradient, with *CXCL9* (G1) and *TNF* (G2) showing complete or partial tolerance and *IL6* and *IL8* showing comparable responsiveness to naive-Mfs (G3) (Figures S5B and S5C). The normal induction of interleukin 6 (IL-6) mRNA expression and the absence of response in ELISA assays therefore suggests that tolerance is a complex phenotype that involves both dampened transcriptional responses to LPS re-exposure and an inability to release some cytokines (Figure S5C). The top tolerance-specific biological process was "cytokine production," while the top pathway was "RIG-I-like signaling" and "p53 signaling" (Figure S5D).

Epigenetic Profile of Tolerized Genes

To understand the molecular mechanisms involved in the altered gene induction by LPS re-exposure in LPS-Mfs, we investigated promoter motif enrichment at overlapping assay for transposase accessible chromatin (ATAC) peaks. Because transcriptional responsiveness to LPS re-exposure in tolerized macrophages occurs on a gradient (Figure 4A), motif enrichment at promoters was scanned in a sliding window of 100 promoters throughout the response gradient from most tolerized (G1) to responsive (G3) genes (Figure 4B). This analysis identified discrete motif signatures in the G1 and G2 tolerized groups. The G1 gene promoters were enriched for several TF motifs, including EGR2, HIF1A, and p53. The latter TF was also identified as a top tolerized pathway (Figure S5D). The partially tolerized genes are enriched for IRF and STAT motifs (Figure 4B). Random Forest analysis also indicated that EGR2 was the top identifier for the G1 group, while IRF and STAT motifs are top identifiers for the G2 group. Interestingly, the G3 group does not contain positive identifiers (Figure 4C). IRF and STAT genes show a tolerized pattern (Figure S6A), indicating that their unresponsiveness to LPS effects downstream partially tolerized genes. On the other hand, *NFKB1* and *RELA* showed normal induction in LPS-Mfs (Figure S6B).

Dynamic H3K27ac change during differentiation and LPS or BG exposure was plotted over the promoter regions of G1, G2, and G3 genes (Figure 4D). Dynamic promoter H3K27ac was observed for roughly half of all genes, with the rest showing consistent high acetylation during all time-points, including LPS re-exposure (not shown). Tolerized genes (G1) and partially tolerized genes (G2) showed no or impaired accumulation of H3K27ac, respectively, after LPS re-exposure in LPS-Mfs compared to naive-Mfs and BG-Mfs (Figure 4D), while responsive genes (G3) were equally acetylated after LPS re-exposure in all subtypes (Figure 4D). H3K4me3 patterns at these promoters closely matched those of H3K27ac (Figure S5B). This finding suggests that LPS-Mfs fail to accumulate H3K27ac at tolerized genes either through absence of pro-inflammatory activators, such as IRF and STATs in the case of G2 genes, or through presence of tolerance inducing TFs, such as HIF1A in the case of G1 genes.

β-Glucan Exposure Can Reverse Tolerance in Both In Vitro and In Vivo LPS-Exposed Monocytes

As indicated before, BG and LPS have an opposing effect on *EGR2* and *MITF* expression (Figure S4A), accumulation of H3K27ac at target enhancers and promoters (Figure 1C), and expression of genes involved in macrophage function, such as lipid metabolism and lysosome and cytokine production (Figure 1E). These findings point to a potential for reversal of LPS-induced tolerance by using BG to stimulate the dectin-1 pathway. To test this hypothesis, monocytes were exposed to LPS for 24 hr and then to BG for 24 hr, followed by a rest period before LPS re-exposure (Figure 5A). We refer to these macrophages as "rescue-Mfs". Additionally, we used the clinically relevant small molecular histone mimic bromodomain and extra-terminal domain family (BET) inhibitor (IBET)151 in a co-treatment with LPS ("preventative") or following LPS exposure ("reversal") setting (Figure 5A). ELISAs showed that BG exposure was able to reverse LPS-induced tolerance and reinstate normal levels of cytokine release in rescue-Mfs (Figure 5B). On the other hand, IBET151 was only effective in preventing tolerance when used to block the LPS-induced response, but it did not reverse tolerance when administered after LPS (Figure 5C). This is in line with the finding that IBET151 is effective in blocking inflammation-associated death in mice (Nicodeme et al., 2010) but suggests that IBET151 is not an effective treatment in monocytes that have already experienced an inflammatory response. Therefore, BG represents a possible treatment option for restoring proper macrophage cytokine release during the post-inflammation tolerance phase.

The suitability of the in vitro tolerance model to mimic the in vivo situation is a major question. Chiefly, does LPS exposure in vivo induce the same transcriptional responses in monocytes, and can in vivo LPS-induced tolerance be reversed by BG? To answer these questions, we used an in vivo experimental human endotoxemia model (Draisma et al., 2009) (Figure 5D). In this model, healthy volunteers are injected with 2 ng/kg US Standard Reference Endotoxin *Escherichia coli* O:113 LPS (Pharmaceutical Development Section of the National Institute of Health, Bethesda, MD, USA), which leads to a sepsis-like state (reviewed in Bahador and Cross, 2007). Study protocols were approved by the local ethics committee of the Radboud University Nijmegen Medical Centre. The volunteers experience transient fever and cold chills as well as pro- and anti-inflammatory cytokine signatures. The in vivo LPS-exposed monocytes show elevated

promoters are E2F3, EGR2, and ZBTB7B motifs. The top features for G2 gene promoters are IRF and STAT, while G3 promoters do not have over-represented features but are depleted of EGR2, E2F3, and ZNF350.
(D) Median H3K27ac at dynamic promoters of G1, G2, and G3 group genes, shaded areas represent the 25th and 75th quartile. This shows that LPS-Mfs do not accumulate H3K27ac at tolerized genes but do so at the promoters of responsive genes. See also Figure S6 for H3K4me3.
See also Figure S6.

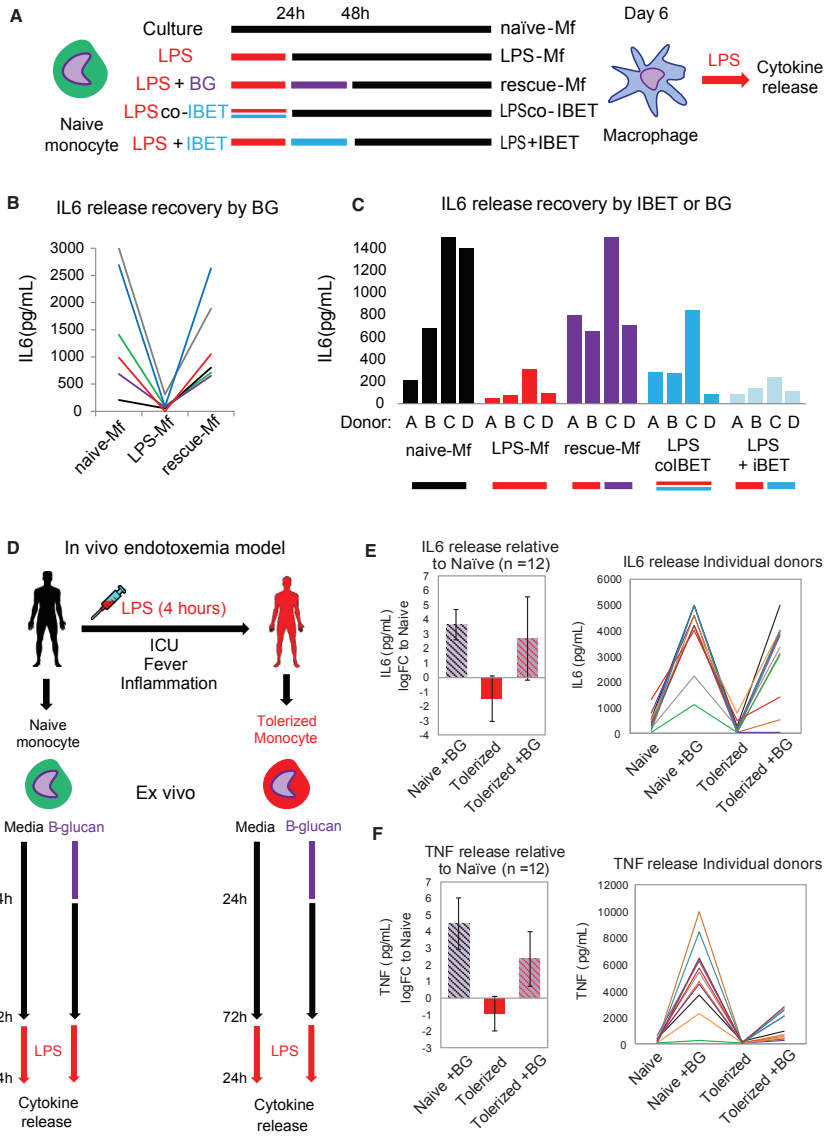


Figure 5. BG Can Reverse Both In Vitro and In Vivo LPS-Induced Tolerance and Reinstate Proper Cytokine Production in Macrophages

(A) The in vitro monocyte tolerance reversal model, with BG added therapeutically after 24 hr of LPS exposure (rescue-Mfs). The histone-mimic and inflammation blocker IBET was used in a preventative (co-culture with LPS for 24 hr LPS-co-IBET-Mfs) and a therapeutic (added after 24 hr of LPS exposure [LPS + IBET-Mfs]) manner. Following several days of rest, macrophages were re-exposed to LPS and cytokine release measured after 24 hr.

(B) BG re-instates IL-6 release in tolerized macrophages. Data from six donors are shown for naive-Mfs, LPS-Mfs, and rescue-Mfs.

(legend continued on next page)

mRNA expression of key cytokines at 4 hr (not shown) and fail to release cytokines in response to a second ex vivo LPS exposure. In this regard, they behave much like in vitro LPS-tolerized monocytes. Monocytes were isolated from peripheral blood taken before and after LPS administration, and then exposed ex vivo to either culture medium alone, or with BG. Cytokine release was measured following LPS re-exposure in culture (Figure 5D). Ex vivo BG exposure increased the release of tumor necrosis factor (TNF) and IL-6 in tolerized monocytes at LPS re-exposure (Figures 5B and 5C). This finding indicates that BG can restore cytokine production of in-vivo-tolerized monocytes. Cumulatively, this confirms that the mechanisms involved in the establishment of tolerance by LPS in vivo and in vitro are similar, validating the use of the in vitro model to study reversal of tolerance by BG. More importantly, it suggests that the BG effect on monocyte tolerance may be transferred to the clinic in the future.

β-Glucan Recovers the Transcriptional Response to LPS at Tolerized Genes

Next, we assessed whether BG reverses tolerance at the transcriptional level. In this experimental setup, monocytes were exposed to LPS followed by BG and then left to rest 24 hr or 4 days before LPS re-exposure for 4 hr (Figure 6A). Additionally, monocytes were treated with a combination of LPS and IBET151 (preventative) and LPS followed by IBET151 (reversal). BG was able to recover the induction of 60% of tolerized genes at day 6 (Figure 6A), including several pro-inflammatory TFs (Figure 6B). Similar effects were observed when IBET151 was used in a preventative model, indicating that BG reversal of LPS-induced tolerance leads to an outcome similar to that produced by blocking LPS-induced tolerance altogether. Overall BG reversal led to a higher median expression of tolerized genes compared to both preventative and reversal use of IBET151 (Figure 6A).

Epigenomic Analysis of β-Glucan Recovery of Tolerized Macrophages

BG exposure following LPS exposure recovers the expression of genes involved in lipid biosynthesis, phagocytosis, and cytokine transport (Figures S7A and S7B). Recovery of expression was observed as early as day 3 and maintained at a higher level in macrophages at day 6 (Figure S7B). Interestingly, addition of BG to naive and tolerized monocytes at day 1, elicited the expression of *EGR2* and *MITF* within 4 hr, with a lower induction in tolerized monocytes (Figure S7C). These findings indicate that BG-induced receptor pathways remain at least partially inducible after the LPS-induced cytokine response and that these pathways can partially recover the naive macrophage epigenetic and transcriptional programs. Analysis of dynamic H3K27ac promoters and enhancers in naive-Mfs, LPS-Mfs, rescue-Mfs, and LPS-co-IBET151-Mfs revealed that BG exposure restores H3K27ac deposition at regions where H3K27ac increase was not obtained following LPS exposure (Figure 7). Interestingly, while IBET151 blocks ~75% of the transcriptional response to LPS in monocytes at 4 hr (data not shown), LPS-co-IBET151-Mfs look more like LPS-Mfs at day 6, indicating no effect of IBET151 on the overall epigenomic profile of LPS-Mfs (Figure 7A, blue square). The effect of BG exposure on H3K27ac deposition in LPS-Mfs was observable at both promoters and distal enhancers of genes involved in metabolism and lipid biosynthesis (Figures 7C and 7D).

DISCUSSION

Perturbation of normal monocyte-to-macrophage differentiation by exogenous signals, such as high bacterial burden in sepsis, can lead to a changed chromatin state and an associated deviation from steady-state function (Amit et al., 2016). This phenomenon is known as innate immune memory, with the best-characterized outcomes being endotoxin tolerance or trained innate immunity (Netea et al., 2016). Trained immunity can have beneficial effects through priming of macrophages for stronger responses to subsequent infection and can be induced by a variety of MAMPs, such as *C. albicans* (Quintin et al., 2012), Bacille Calmette-Guérin (BCG) vaccine (Kleinnijenhuis et al., 2012), and BG (Saeed et al., 2014). Conversely, exposure to high levels of LPS can induce a tolerized macrophage phenotype, which is a major cause of sepsis-associated mortality (SepsisReport, 2012). Previously, we showed that tolerized macrophages (LPS-Mfs) and trained macrophages (BG-Mfs) have distinct epigenetic (Saeed et al., 2014), and metabolic states (Cheng et al., 2014). Mouse studies have shown that such distal element markings are important for appropriate responses to infection (Ghisletti et al., 2010; Ostuni et al., 2013) and identity of tissue-resident macrophages (Amit et al., 2016; Lavin et al., 2014). Nevertheless, until now, the epigenetic basis for endotoxin tolerance in humans has not been explored.

In the current study, our aim was to unveil the early epigenetic and transcriptional events following monocyte exposure to LPS or BG and how the resulting epigenetic landscapes determine the function of tolerized and trained macrophages. LPS- and BG-induced active histone dynamics were observed as early as 1 hr and 4 hr after exposure, respectively (Figures 1B and 1C). Generally, H3K27ac accumulation was accompanied by H3K4me1 accumulation, most obviously at BG-induced enhancers (Figure S1B). Contrary to this general pattern, LPS-induced active enhancers, associated with an inflammation response, showed discordance in time with accumulation of H3K4me1 (Figure S1B), which remained at higher levels in LPS-Mfs compared to naive-Mfs and BG-Mfs. This persistence of H3K4me1 in LPS-Mfs contributes to the overall epigenetic signature of this macrophage subtype and may account for some of the tolerized phenotype (Figure 1B). More pointedly, we discovered a set of more than 3,000 de novo macrophage

(C) Preventative use of IBET blocks the first LPS response in monocytes, resulting in differentiation of macrophages that can release cytokines at the second LPS exposure. Therapeutic use of IBET does not re-instate cytokine release in macrophages.
(D) Experimental human endotoxemia model, with ex vivo BG administration. Monocytes were isolated from 12 healthy volunteers before (naive) and 4 hr after LPS injection (tolerized). Naive or tolerized monocytes were exposed to BG for 24 hr, followed by culture media, or culture media alone. After 3 days ex vivo, monocytes were re-exposed to LPS, and cytokines were measured 24 hr later.
(E and F) BG recovered IL-6 release in 9 out of 12 tolerized monocytes (E) and TNF release in 8 out of 12 monocytes (F). Data are presented as mean ± SD.

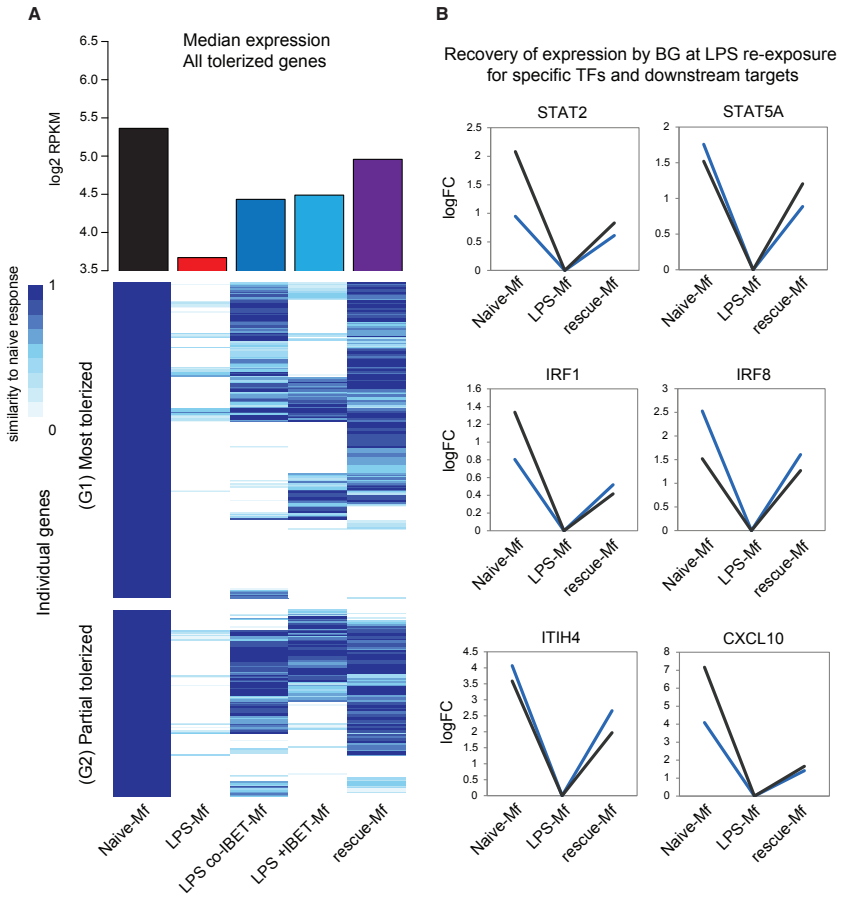


Figure 6. Reversal of Tolerance by BG at the Transcriptional Level

(A) Heatmap of the transcriptional response of naive-Mfs, LPS-Mfs, and rescue-Mfs (BG reversed LPS-Mfs) to LPS re-exposure at day 6. The scale represents relative expression between LPS exposed naive-Mfs (1) and LPS-Mfs (0). Rescue-Mfs exposed to LPS show the most similar profile to naive-Mfs. On the top of the heatmap is median expression (log2 RPKM) of tolerized genes at day 6 and LPS re-exposure in naive-Mfs (black), LPS-Mfs (red), LPS-co-IBET-Mfs (blue), LPS + IBET-Mfs (light blue), and rescue-Mfs (purple).

(B) BG reverses the tolerization of key LPS-induced transcription factors, such as STAT2, STAT5A, IRF1, and IRF8. Log2 fold change increase in mRNA expression is shown.

established distal enhancers that were modulated in the opposite direction by BG or LPS exposure (Figure 1D). Deposition of H3K27ac and H3K4me1 at these regions was accelerated by BG exposure and delayed or completely blocked by LPS exposure. Accordingly, expression of genes near these elements was induced by BG, peaking at 24 hr post-exposure, while they remained lowly expressed in LPS-exposed monocytes (Figure 1F). These genes were involved in lipid metabolism and biosynthesis, phagocytosis, and lysosome maturation (Figures 1G and S2) and have clear TF motif signatures for EGR2, MITF, and ARNT (Figure 2A). Interestingly, *EGR2*, a TF downstream of the BG receptor dectin-1, showed clear transient upregulation by BG but remained inactive in LPS-exposed monocytes, suggesting a possible role in modulating these pathways (Figure S4). TFs and pathways linking lipid biosynthesis and inflammation have been described (Spann et al., 2012). Further, macrophage

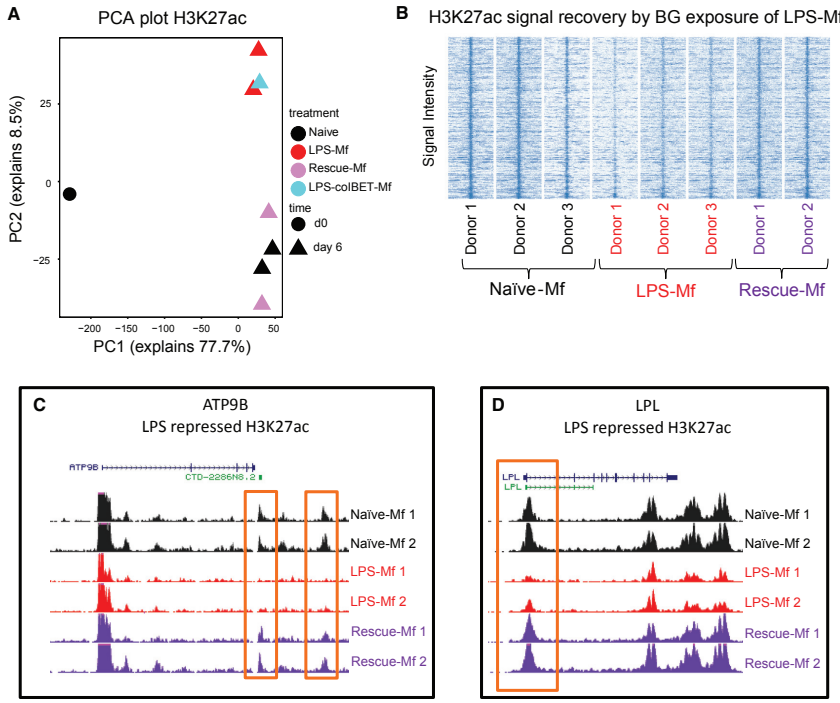


Figure 7. Reversal of Tolerance by BG at the Chromatin Level

(A) PCA plot of H3K27ac dynamics among monocytes, naive-Mfs, LPS-Mfs, rescue-Mfs, and LPS-co-IBET-Mfs. BG exposure of tolerized monocytes results in a H3K27ac profile more similar to naive macrophages, while co-incubation of monocytes with LPS and IBET does not lead to activation of these regions.

(B) Heatmap showing re-establishment of the naive-Mf H3K27ac signal by BG exposure in tolerized macrophages.

(C) PCA plot of H3K4me1 dynamics among monocytes, naive-Mfs, LPS-Mfs, Rescue Mfs, and LPS-co-IBET-Mfs, the effect is similar to H3K27ac, but to a lesser extent. (C) H3K27ac tracks at ATP9B (glucose transport) gene enhancer and (D) LPL (lipid metabolism) gene promoter.

See also Figure S7.

response to infection requires a substantial amount of energy, and shifts in metabolism and energy production are a whole mark of macrophage polarization to M1 or M2 subtypes (Ghesquière et al., 2014), as well as for establishment of trained immunity (Cheng et al., 2014).

Reversal of tolerance after the initial inflammation phase has garnered interest because of the limited success of inflammation-blocking treatments to reduce overall sepsis mortality (Angus and van der Poll, 2013) and because the majority of sepsis deaths occur due to secondary hospital infection during the tolerized phase (Gilroy and Yona, 2015). Our hypothesis was that BG can reverse LPS-induced tolerance because it discordantly regulated pathways that LPS also affected. Specifically, LPS fails to activate key regulators of lipid, lysosome, and metabolism genes, EGR2 and MITF, while BG induces their expression (Figures 1 and 2). Recently, IFNG was shown to partially recover metabolic function in tolerized monocytes from sepsis patients, indicating that reversal of tolerance using innate immune "trainers" is a viable therapeutic strategy (Cheng et al., 2016). We show that BG exposure can indeed reverse the tolerance in macrophages induced by LPS exposure, with rescue-Mfs showing higher release of cytokines in response to a second LPS stimulus (Figures 5A–5C). This was in contrast to the inflammation blocker IBET151, which only prevented tolerance when used to block the initial LPS response but could not reverse it when given to cells after LPS-induced inflammation (Figure 5). In order to further relate our findings to the in vivo situation, we used an experimental human endotoxemia model to induce tolerance in vivo (Draisma et al., 2009; Kox et al., 2014). In terms of cytokine production, in-vivo-tolerized monocytes behave similarly to their in-vitro-tolerized counterparts. The tolerized state of in vivo LPS-exposed monocytes is similar to that of ex-vivo-exposed monocytes and, most importantly, can also be rescued by ex vivo BG exposure (Figures 5D–5F),

indicating that the mechanisms controlling monocyte tolerance in vivo can also be reverted to a more responsive phenotype.

In order to determine the ability of BG to reverse tolerance at the molecular level, we first characterized the transcriptional and epigenetic response to a second LPS exposure in tolerized macrophages (Figures 3 and 4). Studies in mouse sepsis models and human sepsis patients have shown that rather than being inert in response to a second LPS exposure, tolerized macrophages show a shift in the specific pathways that they activate (Foster et al., 2007; Shalova et al., 2015). In line with these studies, we show that LPS-Mfs remodel both H3K27ac and gene expression in response to LPS (Figure 3B). However, the starting point of LPS-Mfs is significantly different from that of naive and BG-Mfs, most clearly for H3K4me1 marked enhancers, suggesting that while activation is occurring, the available enhancer repertoire of these cells is limited or not suited. Our analysis identified a gradient in the LPS-Mf response to LPS, with some genes showing a tolerized pattern (no induction) and others showing a responsive pattern (Figure 3C). The most tolerized gene promoters were enriched for EGR2, HIF1A, and p53 motifs, among many others. A potential role for HIF1A is in agreement with a recent transcriptional analysis in monocytes from sepsis patients (Shalova et al., 2015), while the p53 pathway was a top-ranked tolerized identified by Gene Ontology (GO) analysis (Figure S5D). The strongest enrichment at partially tolerized genes was for the IRF and STAT TF motifs that show strong tolerized expression patterns themselves, ie, *IRF1*, *IRF8*, STAT2, and STAT5A (Figure S6A). *IRF8* and its downstream target, *KLF4*, both of which are important regulators of monocyte differentiation (Kurotaki et al., 2013), show a tolerized profile and enrichment at tolerized gene promoters (Figure 4B). Other, non-TF regulators of tolerance, such as *IRAK3*, *HIF1A*, *SOCS3*, and *IDO1*, are all more highly expressed in LPS-Mfs compared to naive-Mfs and BG-Mfs and have previously been associated with endotoxin-induced tolerance (Bessede et al., 2014; Saeed et al., 2014; Shalova et al., 2015).

Rescue macrophages (BG exposed following LPS exposure) were able to induce ~60% of tolerized genes at LPS re-exposure (Figure 6). This indicates that BG reversal of tolerance at the transcriptional level is not complete (Figure 6A). Fascinatingly, BG recovered the expression of tolerized genes to a level greater than that observed in macrophages treated with IBET and LPS together (Figure 6A). This indicates that BG can reinstate a responsive state at a higher level than that obtained by actually blocking the initial LPS transcriptional response. This important observation suggests that BG-associated pathways remain intact even after large-scale epigenetic and transcriptional programs are induced by LPS. At the level of histone modifications, BG recovers H3K27ac at regions that are silent in LPS-Mfs, further supporting the notion that the molecular mechanisms required for BG-induced chromatin remodeling remain after the initial LPS response (Figures 7 and S7).

In conclusion, the hypothesis-free epigenomic and transcriptomic analysis of monocyte-to-macrophage differentiation and innate immune memory generated a number of testable hypotheses. Our findings show that the innate immune "training stimulus" β-glucan can reverse macrophage tolerance ex vivo. This is an important step toward understanding how the tolerized phenotype can be reversed in sepsis patients and ultimately provides the framework for future therapeutic developments in innate immune diseases.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Monocytes from Healthy Donors
  - In Vitro Monocyte-to-Macrophage Differentiation and Induction of Innate Immune Memory
  - Experimental Human Endotoxemia Model
- METHOD DETAILS
  - Cytokine Assays
  - RNA Extraction and cDNA Synthesis
  - Chromatin Immunoprecipitation
  - Library Preparation for Sequencing
  - Assay for Transposase Accessible Chromatin
  - Whole Genome Bisulfite Sequencing
  - RNA-Seq Data Analysis
  - ChIP-Seq Data Analysis
  - ATAC-Seq Data Analysis
  - DNA-Binding Motif Scanning
  - Gene Ontology Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY
  - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2016.09.034.

A video abstract is available at http://dx.doi.org/10.1016/j.cell.2016.09.034#mmc6.

AUTHOR CONTRIBUTIONS

Conceptualization, H.G.S., C.L., J.H.A.M., and M.G.N.; Methodology, B.N., R.D., B.K., R.J.W.A., T.K., F.M., C.W., E.M.J.-M., J.Z., M.K., P.P., and N.S.; Investigation, B.N., E.H., S.-Y.W., and W.M.; Writing – Original Draft, B.N., H.G.S., and C.L.; Writing – Review & Editing, E.H., S.-Y.W., M.G.N., J.H.A.M., R.J.W.A., and T.K.; Funding Acquisition, H.G.S.; Resources, P.F., L.C., S.J.v.H., P.P., and I.G.; Supervision, H.G.S., C.L., J.H.A.M., and M.G.N.

ACKNOWLEDGMENTS

The research leading to the results described in this paper has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement 282510-BLUEPRINT. B.N. is supported by an NHMRC (Australia) CJ Martin Early Career Fellowship. M.G.N. was supported by an ERC Consolidator Grant (310372). The authors would like to thank GSK Epinova and Cellzome for providing the IBET reagent and Prof. David L. Williams (University of Tennessee) for β1,3(D)glucan (β-glucan).

Received: January 15, 2016
Revised: August 5, 2016
Accepted: September 20, 2016
Published: November 17, 2016

REFERENCES

Amit, I., Winter, D.R., and Jung, S. (2016). The role of the local environment and epigenetics in shaping macrophage identity and their effect on tissue homeostasis. Nat. Immunol. *17*, 18–25.

Angus, D.C., and van der Poll, T. (2013). Severe sepsis and septic shock. N Engl J Med. *369*, 840–851.

Bahador, M., and Cross, A.S. (2007). From therapy to experimental model: a hundred years of endotoxin administration to human subjects. J. Endotoxin Res. *13*, 251–279.

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics *27*, 1691–1692.

Bessede, A., Gargaro, M., Pallotta, M.T., Matino, D., Servillo, G., Brunacci, C., Bicciato, S., Mazza, E.M., Macchiarulo, A., Vacca, C., et al. (2014). Aryl hydrocarbon receptor control of a disease tolerance defence pathway. Nature *511*, 184–190.

Biswas, S.K., and Lopez-Collazo, E. (2009). Endotoxin tolerance: new mechanisms, molecules and clinical significance. Trends Immunol. *30*, 475–487.

Cheng, S.C., Quintin, J., Cramer, R.A., Shepardson, K.M., Saeed, S., Kumar, V., Giamarellos-Bourboulis, E.J., Martens, J.H., Rao, N.A., Aghajanirefah, A., et al. (2014). mTOR- and HIF-1α-mediated aerobic glycolysis as metabolic basis for trained immunity. Science *345*, 1250684.

Cheng, S.C., Scicluna, B.P., Arts, R.J., Gresnigt, M.S., Lachmandas, E., Giamarellos-Bourboulis, E.J., Kox, M., Manjeri, G.R., Wagenaars, J.A., Cremer, O.L., et al. (2016). Broad defects in the energy metabolism of leukocytes underlie immunoparalysis in sepsis. Nat. Immunol. *17*, 406–413.

de la Rica, L., Rodríguez-Ubreva, J., García, M., Islam, A.B., Urquiza, J.M., Hernando, H., Christensen, J., Helin, K., Gómez-Vaquero, C., and Ballestar, E. (2013). PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. Genome Biol. *14*, R99.

Draisma, A., Pickkers, P., Bouw, M.P., and van der Hoeven, J.G. (2009). Development of endotoxin tolerance in humans in vivo. Crit. Care Med. *37*, 1261–1267.

Foster, S.L., Hargreaves, D.C., and Medzhitov, R. (2007). Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature *447*, 972–978.

Ghesquière, B., Wong, B.W., Kuchnio, A., and Carmeliet, P. (2014). Metabolism of stromal and immune cells in health and disease. Nature *511*, 167–176.

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L., et al. (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity *32*, 317–328.

Gilroy, D.W., and Yona, S. (2015). HIF1α allows monocytes to take a breather during sepsis. Immunity *42*, 397–399.

Glass, C.K., and Natoli, G. (2016). Molecular control of activation and priming in macrophages. Nat. Immunol. *17*, 26–33.

Goodridge, H.S., Simmons, R.M., and Underhill, D.M. (2007). Dectin-1 stimulation by Candida albicans yeast or zymosan triggers NFAT activation in macrophages and dendritic cells. J. Immunol. *178*, 3107–3115.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kleinnijenhuis, J., Quintin, J., Preijers, F., Joosten, L.A., Ifrim, D.C., Saeed, S., Jacobs, C., van Loenhout, J., de Jong, D., Stunnenberg, H.G., et al. (2012). Bacille Calmette-Guerin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. Proc. Natl. Acad. Sci. USA *109*, 17537–17542.

Kox, M., van Eijk, L.T., Zwaag, J., van den Wildenberg, J., Sweep, F.C., van der Hoeven, J.G., and Pickkers, P. (2014). Voluntary activation of the sympathetic nervous system and attenuation of the innate immune response in humans. Proc. Natl. Acad. Sci. USA *111*, 7379–7384.

Kuhn, M. (2008). Building predictive models in R using the caret package. J Stat Softw. *28*, 1–26.

Kulis, M., Merkel, A., Heath, S., Queirós, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., et al. (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. Nat. Genet. *47*, 746–756.

Kurotaki, D., Osato, N., Nishiyama, A., Yamamoto, M., Ban, T., Sato, H., Nakabayashi, J., Umehara, M., Miyake, N., Matsumoto, N., et al. (2013). Essential role of the IRF8-KLF4 transcription factor cascade in murine monocyte differentiation. Blood *121*, 1839–1849.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell *159*, 1312–1326.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Mammana, A., and Chung, H.R. (2015). Chromatin segmentation based on a probabilistic model for read counts explains a large portion of the epigenome. Genome Biol. *16*, 151.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Netea, M.G., Joosten, L.A., Latz, E., Mills, K.H., Natoli, G., Stunnenberg, H.G., O'Neill, L.A., and Xavier, R.J. (2016). Trained immunity: A program of innate immune memory in health and disease. Science *352*, aaf1098.

Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C.W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., et al. (2010). Suppression of inflammation by a synthetic histone mimic. Nature *468*, 1119–1123.

Nishikawa, K., Iwamoto, Y., Kobayashi, Y., Katsuoka, F., Kawaguchi, S., Tsujita, T., Nakamura, T., Kato, S., Yamamoto, M., Takayanagi, H., and Ishii, M. (2015). DNA methyltransferase 3a regulates osteoclast differentiation by coupling to an S-adenosylmethionine-producing metabolic pathway. Nat. Med. *21*, 281–287.

Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. Cell *152*, 157–171.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Quintin, J., Saeed, S., Martens, J.H., Giamarellos-Bourboulis, E.J., Ifrim, D.C., Logie, C., Jacobs, L., Jansen, T., Kullberg, B.J., Wijmenga, C., et al. (2012). Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. Cell Host Microbe *12*, 223–232.

Quintin, J., Cheng, S.C., van der Meer, J.W., and Netea, M.G. (2014). Innate immune memory: towards a better understanding of host defense mechanisms. Curr. Opin. Immunol. *29*, 1–7.

Rialdi, A., Campisi, L., Zhao, N., Lagda, A.C., Pietzsch, C., Ho, J.S., Martinez-Gil, L., Fenouil, R., Chen, X., Edwards, M., et al. (2016). Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation. Science *352*, aad7993.

Saeed, S., Quintin, J., Kerstens, H.H., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.C., Ratter, J., Berentsen, K., van der Ent, M.A., et al. (2014). Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science *345*, 1251086.

SepsisReport (2012). Focus on sepsis. Nat. Med. *18*, 997.

Shalova, I.N., Lim, J.Y., Chittezhath, M., Zinkernagel, A.S., Beasley, F., Hernández-Jiménez, E., Toledano, V., Cubillos-Zapata, C., Rapisarda, A., Chen, J., et al. (2015). Human monocytes undergo functional re-programming during sepsis mediated by hypoxia-inducible factor-1α. Immunity *42*, 484–498.

Spann, N.J., Garmire, L.X., McDonald, J.G., Myers, D.S., Milne, S.B., Shibata, N., Reichart, D., Fox, J.N., Shaked, I., Heudobler, D., et al. (2012). Regulated accumulation of desmosterol integrates macrophage lipid metabolism and inflammatory responses. Cell *151*, 138–152.

van Heeringen, S.J., and Veenstra, G.J. (2011). GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. Bioinformatics *27*, 270–271.

Vento-Tormo, R., Company, C., Rodríguez-Ubreva, J., de la Rica, L., Urquiza, J.M., Javierre, B.M., Sabarinathan, R., Luque, A., Esteller, M., Aran, J.M., et al. (2016). IL-4 orchestrates STAT6-mediated DNA demethylation leading to dendritic cell differentiation. Genome Biol. *17*, 4.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431–1443.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics *26*, 873–881.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-H3K27ac	Diagenode	pAb-196-050
Rabbit polyclonal anti-H3K4me1	Diagenode	pAb-037-050
Rabbit polyclonal anti-H3K4me3	Diagenode	pAb-003-050
Rabbit polyclonal anti-H3K27me3	Diagenode	pAb-195-050
Rabbit polyclonal anti-H3K9me3	Diagenode	pAb-193-050
Chemicals, Peptides, and Recombinant Proteins		
IBET-151	GSK Epinova and Cellzome	GSK1210151A
Human Serum	Sigma-Aldrich	H4522-100ML
RPMI 1640 Medium, GlutaMAX	Thermo Fisher Scientific	61870036
Gentamycin	Thermo Fisher Scientific	15750060
L-glutamine	Thermo Fisher Scientific	25030081
Sodium Pyruvate	Thermo Fisher Scientific	11360070
Percoll	Sigma-Aldrich	P1644-1L
Ficoll Paque Plus	Sigma-Aldrich	GE17-1440-03
Lipopolysaccharides from *Escherichia coli* 055:B5	Sigma-Aldrich	L2880-10MG
β1,3(D)glucan (β-glucan)	(Saeed et al., 2014)	N/A
2-Mercaptoethanol	Thermo Fisher Scientific	21985023
Actinomycin D	Thermo Fisher Scientific	11805017
IGEPAL CA-630	Sigma-Aldrich	I8896-50ML
Critical Commercial Assays		
KAPA library preparation kit	Kapa Biosystems	KK8400
riboZero gold rRNA removal kit	Illumina	MRZG12324
Nextera DNA Library Prep Kit	Illumina	FC-121-1031
TruSeq SBS KIT v3 - HS (50 cycles)	Illumina	FC-401-3002
NextSeq 500/550 High Output v2 kit (75 cycles)	Illumina	FC-404-2005
NEBNext High-Fidelity 2 × PCR Master Mix	New England Biolabs	M0541
iQ SYBR Green Supermix	Bio-Rad	1708880
100 × SYBR Green I Nucleic Acid Gel Stain	Thermo Fisher Scientific	S7563
Human IL-6 elisa	Sanquin	M9316
Human TNFα elisa	R&D	DY210
SPRIselect reagent kit	Beckman Coulter	B23218
E-Gel SizeSelect Agarose Gels, 2%	Thermo Fisher Scientific	G661002
CD3 MicroBeads, human	Miltenyi Biotec	130-050-101
CD19 MicroBeads, human	Miltenyi Biotec	130-050-301
CD56 MicroBeads, human	Miltenyi Biotec	130-050-401
dNTP set 100 mM	Life Technologies	10297-018
dUTP 100 mM	Promega	U119A
Glycogen (20 mg/ml)	Life Technologies	10814-010
Random Hexamer primers	Sigma-Aldrich	11034731001
Second Strand Buffer	Life Technologies	10812-014
Superscript III Reverse Transcriptase	Life Technologies	18080-044
DNA polymerase I, *E. coli*	New England Biolabs	M0209S
USER enzyme	New England Biolabs	M5505L

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
E.Coli Ligase	New England Biolabs	M0205L
Rnasin Plus Rnase Inhibitor	Promega	N2615
Ribonuclease H	Life Technologies	AM2293
T4 DNA polymerase	New England Biolabs	M0203L
Sodium Acetate (3M)	Life Technologies	AM9740
DNase I	QIAGEN	79254
Qubit RNA HS assay kit	Life Technologies	Q32852
Ribozero Gold Kit	Illumina	MRZG12324
Rneasy Mini Kit	QIAGEN	74106
Deposited Data		
Raw data files for RNA sequencing	This paper	GEO: GSE85243
Raw data files for ChIP sequencing	This paper	GEO: GSE85245
Raw data files for ATAC sequencing	This paper	GEO: GSE87218
Raw data files for WGBS sequencing	This paper	EGA: EGAD00001002693
Experimental Models: Organisms/Strains		
Human: primary monocytes from healthy volunteers	Sanquin Blood Bank	N/A
Sequence-Based Reagents		
NEXTflex DNA Barcodes - 48	Bioo Scientific	514104
Primer EGR2: F 5′ TTGACCAGATGAACGGAGTG 3′R 5′ GTTGAAGCTGGGGAAGTGAC 3′	This paper	N/A
Primer MITF: F 5′ AACTCATGCGTGAGCAGATG 3′R 5′ TACTTGGTGGGGTTTTCGAG 3′	This paper	N/A
Primer CSF1: F 5′ CAGATGGAGACCTCGTGCC 3′R 5′ GCATTGGGGGTGTTATCTCTG 3′	This paper	N/A
Primer LAMP1: F 5′ TGAACAAGACAGGCCT TCCC 3′R 5′ TGTGCAGCTCCAGAGTCACC 3′	This paper	N/A
Software and Algorithms		
Bedtools	(Quinlan and Hall, 2010)	http://bedtools.readthedocs.io/en/latest/
Bamtools	(Barnett et al., 2011)	https://github.com/pezmaster31/bamtools
Samtools	(Li and Durbin, 2009)	http://samtools.sourceforge.net/
GSNAP	(Wu and Nacu, 2010)	http://research-pub.gene.com/gmap/
GimmeMotifs	(van Heeringen and Veenstra, 2011)	https://github.com/simonvh/gimmemotifs
Caret	(Kuhn, 2008)	http://cran.r-project.org/web/packages/caret/index.html
HOMER	(Heinz et al., 2010)	http://homer.salk.edu/homer/motif/
DAVID	(Huang da et al., 2009)	https://david.ncifcrf.gov/
bwa	(Li and Durbin, 2009)	http://bio-bwa.sourceforge.net/
bowtie	(Langmead et al., 2009)	http://bowtie-bio.sourceforge.net/index.shtml
MACS2	(Zhang et al., 2008)	https://github.com/taoliu/MACS

CONTACT FOR REAGENTS AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by the corresponding author Hendrik G. Stunnenberg (h.stunnenberg@ncmls.ru.nl).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Monocytes from Healthy Donors

All primary cells were isolated from healthy volunteers who gave written informed consent (Sanquin Blood bank, Nijmegen, the Netherlands). Volunteers are of Northern European descent. Peripheral blood mononuclear cells were isolated by centrifugation in

Ficoll-Paque (GE Healthcare), followed by removal of T cells using an additional Percoll gradient. Monocytes were purified from PBMCs using negative selection in an LD column magnet separator, with beads for CD3+ (T cells), CD19+ (B cells) and CD56+ (NK cells) positive cells (Miltenyi Biotech), yielding > 95% pure monocytes. Successful isolation of monocytes was confirmed with FACS, as previously described (Saeed et al., 2014).

In Vitro Monocyte-to-Macrophage Differentiation and Induction of Innate Immune Memory

Monocytes were differentiated into resting macrophages by ex vivo culture in RPMI 1640 medium (Sigma Aldrich) with 10% Human Serum. Media was supplemented with 10 μg/mL gentamycin, 10 mM L-glutamine and 10 mM pyruvate (Life Technologies). Tolerization was induced by treatment of monocytes with 10-100ng/mL LPS for 24 hr, followed by washout and five days culture in RPMI + 10% human serum, while trained innate immunity was induced by treatment with 5 μg/mL BG for 24 hr, followed by washout and 5 days in culture. Establishment of tolerance or training in the resulting macrophages at day 6 was determined by TNF and IL6 release at 24 hr following LPS stimulation using ELISA. For ChIP-seq, $10x10^6$ monocytes were seeded in 10cm dishes, for RNaseq and ATAC-seq 1.5×10^6 monocytes were seeded in 6 well plates. IBET151 (GSK) was diluted to 50 mM stock using DMSO. Following dosage titration 5 μM was determined as the appropriate final concentrations to prevent tolerization, without causing cell death. IBET-151 was added to monocytes at the same time as LPS for 24 hr, followed by washout and five days culture in RPMI + 10% human serum to macrophage differentiation.

Experimental Human Endotoxemia Model

In vivo endotoxin tolerance was examined in 12 healthy nonsmoking volunteers who participated in an experimental human endotoxemia study. The study is registered at Clinicaltrials.gov (NCT02602977) and study protocols were approved by the local ethics committee of the Radboud University Nijmegen Medical Centre (NL53584.091.15/CMO 2015-1796). Written informed consent was obtained from all study participants. Subjects were screened before the start of the experiment and had a normal physical examination, electrocardiography, and routine laboratory values. Throughout the study period, subjects were not allowed to take any drugs, including acetaminophen, and were asked to refrain from alcohol and caffeine 24 hr and from food 12 hr before the start of the endotoxemia experiment. All study procedures were conducted in accordance with the declaration of Helsinki including current revisions and Good Clinical Practice guidelines. Experimental human endotoxemia was conducted as described previously (Kox et al., 2014). Briefly, all subjects received an intravenous bolus injection of LPS (lipopolysaccharide derived from *Escherichia coli* O:113, Clinical Center Reference Endotoxin, National Institutes of Health (NIH), Bethesda, MD) at a dose of 2 ng/kg. Blood was obtained before LPS administration and 4 hr afterward, and monocytes were isolated. Monocytes were exposed to culture or BG ex vivo, and cytokine production in the supernatants was measured following ex vivo LPS (10ng/ml) exposure. Cytokine production was determined by ELISA following the protocol of the manufactures (IL-6, sanquin and TNFα, R&D systems).

METHOD DETAILS

Cytokine Assays

TNFα and IL-6 were measured using ELISA according to the manufacturer protocol (IL6: Sanquin; and TNFα: R&D). For cytokines production assays the differences between groups were analyzed using the Wilcoxon signed-rank test. The level of significance was defined as a p value < 0.05.

RNA Extraction and cDNA Synthesis

Total RNA was extracted from cells using the QIAGEN RNeasy RNA extraction kit (QIAGEN, Netherlands), using on-column DNaseI treatment. Ribosomal RNA was removed using the riboZero rRNA removal kit (Illumina). RNA was then fragmented into 200bp fragments by incubation for 7.5 min at 95°C in fragmentation buffer (200 mM Tris-acetate, 500 mM Potassium Acetate, 150 mM Magnesium Acetate [pH 8.2]). First strand cDNA synthesis was performed using SuperScript III (Life Technologies), followed by synthesis of the second cDNA strand. Library preparation was performed using the KAPA hyperprep kit (KAPA Biosystems). Quality of cDNA and the efficiency of ribosomal RNA removal was confirmed using quantitative RT-PCR using the IQ Sybr Supermix, with primers for GAPDH, 18S and 28S rRNA.

Chromatin Immunoprecipitation

Purified cells were fixed with 1% formaldehyde (Sigma) at a concentration of approximately 10 million cells/ml. Fixed cell preparations were sonicated using a Diagenode Bioruptor UCD-300 for 3x 10 min (30 s on; 30 s off). 67 μl of chromatin (1 million cells) was incubated with 229 μl dilution buffer, 3 μl protease inhibitor cocktail and 0.5-1μg of H3K27ac, H3K4me3, H3K4me1, H3K27me3, H3K9me3 or H3K36me3 antibodies (Diagenode) and incubated overnight at 4°C with rotation. Protein A/G magnetic beads were washed in dilution buffer with 0.15% SDS and 0.1% BSA, added to the chromatin/antibody mix and rotated for 60 min at 4°C. Beads were washed with 400μl buffer for 5 min at 4°C with five rounds of washes. After washing chromatin was eluted using elution buffer for 20 min. Supernatant was collected, 8 μl 5M NaCl, 3μl proteinase K were added and samples were incubated for 4 hr at 65°C.Finally samples were purified using QIAGEN; Qiaquick MinElute PCR purification Kit and eluted in 20 μl EB. Detailed protocols can be found on the Blueprint website (http://www.blueprint-epigenome.eu/UserFiles/file/Protocols/Histone_ChIP_May2013.pdf).

Library Preparation for Sequencing

Illumina library preparation was done using the Kapa Hyper Prep Kit. For end repair and A-tailing double stranded DNA was incubated with end repair and A-tailing buffer and enzyme and incubated first for 30 min at 20°C and then for 30 min at 65°C.Subsequently adapters were ligated by adding 30μl ligation buffer, 10 Kapa I DNA ligase, 5 μl diluted adaptor in a total volume of 110μl and incubated for 15 min at 15°C. Post-ligation cleanup was performed using Agencourt AMPure XP reagent and products were eluted in 20 μl elution buffer. Libraries were amplified by adding 25 μl 2x KAPA HiFi Hotstart ReadyMix and 5μl 10x Library Amplification Primer Mix and PCR, 10 cycles. Samples were purified using the QIAquick MinElute PCR purification kit and 300bp fragments selected using E-gel. Correct size selection was confirmed by BioAnalyzer analysis. Sequencing was performed using Illumina HiSeq 2000 machines and generated 43bp single end reads. Samples for RNA-seq were treated to the above protocol exactly, except for a single additional step: After post-ligation cleanup, and before library amplification, samples were incubated with 3 uL USER enzyme for 15 min at 37°C to digest the 2nd cDNA strand.

Assay for Transposase Accessible Chromatin

Monocytes or macrophages (100,000 cells) were scrapped in a well of a 6-well plate with cold PBS and then spun down at 800 × g for 5 min at 4°C. Cells were washed with 50 μl of cold 1x PBS buffer, incubated in 50 μl of cold lysis buffer (10 mM Tris-HCL (pH 7.4), 10 mM NaCl, 3 mM MgCl2 0, 1% IGEPAL) and spun down at 800 × g for 10 min at 4°C. The nuclei were immediately resuspended in the transposition reaction mix (22.5 μl TD buffer, 2.5 μl Tn5 Transposase, 25 μl NF H2O) and incubated for 30 min at 37°C. Following transposition, 100 μl AMPure beads were added to the reaction (sample-to-bead ratio of 1:2), mixed thoroughly by pipetting, and incubated for 15 min at RT. Samples and beads were washed on the magnetic rack with 80% ethanol, dried for 5 min, and resuspended in 15 μl EB buffer. DNA was amplified with 10 - 15 PCR cycles using the mix (15 μl transposed DNA, 0.3 μl 100x SYBR Green I, 25 μl NEBNext High-Fidelity master mix, 2.5 μl Nextera Primer index N7.. (25 μM), 2.5 μl Nextera Primer index S5.. (25 μM), 4.7 μl NF H2O). In order to reduce GC and size bias in PCR, the PCR reaction is monitored using qPCR to stop amplification prior to saturation. Following amplification, samples were incubated purified twice using SPRI beads, first using negative selection with a sample-to-bead ratio of 1-0.65 and then positive selection with a sample-to-bead ratio of 1-1.8. After 80% Ethanol wash and drying, the sample was eluted in 20 μl EB buffer, and quality checked before sequencing. Detailed protocol can be found on the Blueprint website (http://www.blueprint-epigenome.eu/UserFiles/file/Protocols/ATAC_Seq_Protocol.pdf).

Whole Genome Bisulfite Sequencing

Genomic DNA (1-2 μg) was spiked with unmethylated λ DNA (5ng of λ DNA per μg of genomic DNA) (Promega). The DNA was sheared by sonication to 50-500bp using a Covaris E220 and fragments of size 150-300 bp were selected using AMPure XP beads (Agencourt Bioscience). Genomic DNA libraries were constructed using the Illumina TruSeq Sample Preparation kit (Illumina) following the Illumina standard protocol: end repair was performed on the DNA fragments, an adenine was added to the 3′ extremities of the fragments and Illumina TruSeq adapters were ligated at each extremity. Adter adaptor ligation, the DNA was treated with sodium bisulfite using the EpiTexy Bisulfite kit (QIAGEN) following the manufacturer's instructions for formalin-fixed and paraffin-embedded (FFPE) tissue samples. Two rounds of bisulfite conversion were performed to assure a high conversion rate. An enrichment for adaptor-ligated DNA was carried out through 7 PCR cycles using the PfuTurboCx Hotstart DNA polymerase (Stratagene). Library quality was monitored using the Agilent 2100 BioAnalyzer (Agilent), and the concentration of viable sequencing fragments (molecules carrying adaptors at both extremities) estimated using quantitative PCR with the library quantification kit from KAPA Biosystem. Paired-end DNA sequencing (2x100 nucleotides) was then performed using the Illumina Hi-Seq 2000. WGBS data are available upon request from the BLUEPRINT consortium.

RNA-Seq Data Analysis

For quality control and visualization, RNA-seq reads were aligned to the hg19 reference genome using GSNAP (Wu and Nacu, 2010) with non-default parameters -m 1 -N 1 -n 1 -Q -s Ensembl_splice_68. Each RNA-seq sample was subjected to a quality control step, where, based on read distribution over the annotated genome, libraries that are outliers were identified and discarded from further analysis. To infer gene expression levels, RNA-seq reads were aligned to the Ensembl v68 human transcriptome using Bowtie. Quantification of gene expression was performed using MMSEQ. Differential expression was determined using MMDIFF. A two model comparison was used to identify differentially expressed genes that confer cellular identity Mo/Mf. The null-model is that the mean expression levels are the same in both cell types, and the alternative model is that the mean expression levels are allowed to differ between the two cell types. Genes with a larger posterior probability for the second model, an RPKM value greater than 2 in any of Mo or Mf and minimally a 2-fold expression change were considered as differentially expressed. Expression changes related to differentiation of each treatment were studied using a 52-model comparison, a.k.a. polytomous comparison, under the null-model that assumes the mean expression levels are the same across each time-point. Expression differences related to the treatments at each time-point were studied using a 5-model comparison, under the null-model that assumes the mean expression levels are the same across each treatment.

ChIP-Seq Data Analysis

Sequencing reads were aligned to human genome assembly hg19 (NCBI version 37) using bwa. Duplicate reads were removed after the alignment with the Picard tools. For peak calling the BAM files were first filtered to remove the reads with mapping quality less

than 15, followed by fragment size modeling (https://code.google.com/archive/p/phantompeakqualtools/). MACS2 (https://github.com/taoliu/MACS/) was used to call the peaks. H3K4me1,H3K9me3 and H3K27me3 peaks were called using the broad setting of MACS2 while H3K27ac and H3K4me3 were called using the default (narrow) setting. For each histone mark dataset, the data were normalized using the R package DESeq2 and then pair-wise comparisons were performed (fold change 3, adjusted p-adjvalue < 0.05 and RPKM $\geq$ 2 in at least in any condition) to determine the differentially expressed genes per condition. The results from all possible pairwise comparisons (within each condition and similar time points across all conditions per mark) were pooled and merged to define the dynamic set of enriched regions. Promoters were defined as regions between ± 2kb from TSSs for each ensemble gene and enhancers were determined as enriched H3K27ac/H3K4me1 regions more than ± 2kb away from the TSS. To find different patterns over dynamic promoters or enhancers, we applied *a K-means* clustering procedure (with optimal number of clusters per each dataset) to the dynamic datasets as described above.

ATAC-Seq Data Analysis

The full ATAC-seq protocol is available at the BLUEPRINT website (http://www.blueprint-epigenome.eu/UserFiles/file/Protocols/ATAC_Seq_Protocol.pdf). ATAC-seq reads were mapped to the hg19 reference genome using BWA (Li and Durbin, 2009) with default parameters. Non-uniquely mapped reads and PCR duplicates were removed. MACS2 (Zhang et al., 2008) was used to identify regions of open chromatin (peaks) with parameters "–nomodel -p 1e-9." Overlap peaks from different samples were merged.

DNA-Binding Motif Scanning

All the DNA-binding motifs used in this study are based on the cis-bp database described in (Weirauch et al., 2014). Only motifs with direct evidence of binding in the species of vertebrate were selected. Within each motif family, as annotated by cis-bp, all motifs were clustered using 'gimme cluster' from the GimmeMotifs package (van Heeringen and Veenstra, 2011) with a threshold of 0.9999. The annotation of motifs is based on the annotation of human in the cis-bp database. Motifs were used for scanning if the assigned TF is expressed (> 1 RPKM) in at least one time-point during the differentiation. Total ATAC-seq peaks were scanned for the presence of motifs. We used Gimme motifs for scanning with dynamic motif scoring cut-offs targeting a false discovery rate (FDR) of both 0.01 and 0.05. To look at the motif enrichments in each set of regions (epigenetic cluster or gene cluster), ATAC-seq peaks were assigned to the epigenomic cluster or the gene promoters by intersection. Motif occurrences were acquired by intersection of the assigned ATAC-seq peaks with the motif scanning results on total ATAC-seq peaks. Total ATAC-seq peaks were divided into promoter set and non-promoter set as the background for the calculation of motif enrichment. Enrichment of motifs in each set of regions was defined by applying a hypergeometric test using the motif frequency in the corresponding background. This results in TFs that putatively regulate the activities of the regulatory regions. Motifs in each heat map satisfy an arbitrary cutoff of > 5% motif presence and a fixed minimal presence difference from background in at least one cluster. Hierarchical clustering (Pearson correlation) was performed in each heat map using the motif occurrence frequencies in the clusters. Based on the gene activity and dynamics, only one TF was selected to represent a motif if multiple genes are assigned to the same motif. Scanning results from FDR of 0.01 and 0.05 were compared and do not affect the result of enrichment analysis.

Gene Ontology Analysis

Gene ontology analysis on dynamic lists of genes was performed using DAVID (Huang da et al., 2009). Gene ontology on dynamic enhancer clusters was performed using GREAT (McLean et al., 2010). KEGG pathways and Biological Processes were ranked by p value and the top terms were plotted.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical parameters including the exact value of n, the definition of center, dispersion, and precision measures (mean ± SEM) and statistical significance are reported in the Figures and the Figure Legends. Data are judged to be statistically significant when $p < 0.05$ by two-tailed Student's T-Test or 2-way ANOVA, where appropriate.

DATA AND SOFTWARE AVAILABILITY

Data Resources

Raw data files for the RNA, ATAC, and ChIP sequencing and analysis have been deposited in the NCBI Gene Expression Omnibus under accession number: GSE85246.

Links to GEO SubSeries linked to GSE85246:

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85243
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85245
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87218

Cell

4

Supplemental Figures

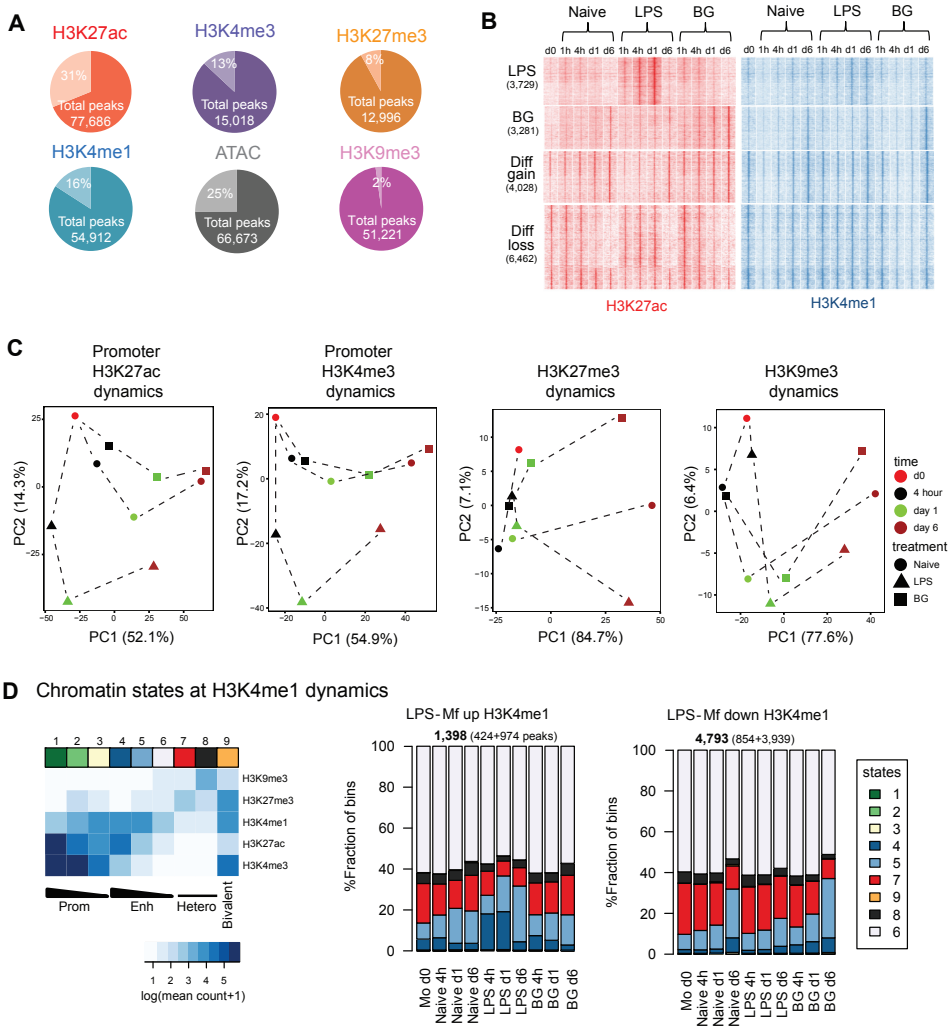


Figure S1. Summary of Dynamic Histone Marks and PCA Plots of Dynamic Active Histone Modifications at Promoters and Repressive Marks, Related to Figure 1

(A) Percentage of histone ChIP-seq peaks designated as dynamic across time-points and between treatments. H3K27ac was the most dynamic modification, with almost a third of regions showing significant changes.

(B) Heatmap showing histone intensity of H3K27ac and H3K4me1 at dynamic H3K27ac enhancers with 12kb ± from center of the peak.

(C) PCA plots for all time-points for H3K27ac dynamic promoters, H3K4me3 dynamic promoters, dynamic H3K27me3 regions, and dynamic H3K9me3 regions. H3K27ac and H3K4me3 at promoters behave similarly over time and in response to LPS or BG exposure, and reflect the behavior of H3K27ac at enhancers. Unlike active marks, repressive marks show little dynamics up to day 1.

(D) Stacked plots showing chromatin state changes over differentiation at "LPS-Mf up" and "BG up / LPS down" H3K4me1 enhancers. These enhancers are established through H3K27ac dynamics shown in Figure 1C. The genome was segmented into 9 chromatin states based on the 5 histone marks analyzed. This analysis indicates that H3K4me1 increase is associated with loss of H3K27me3.

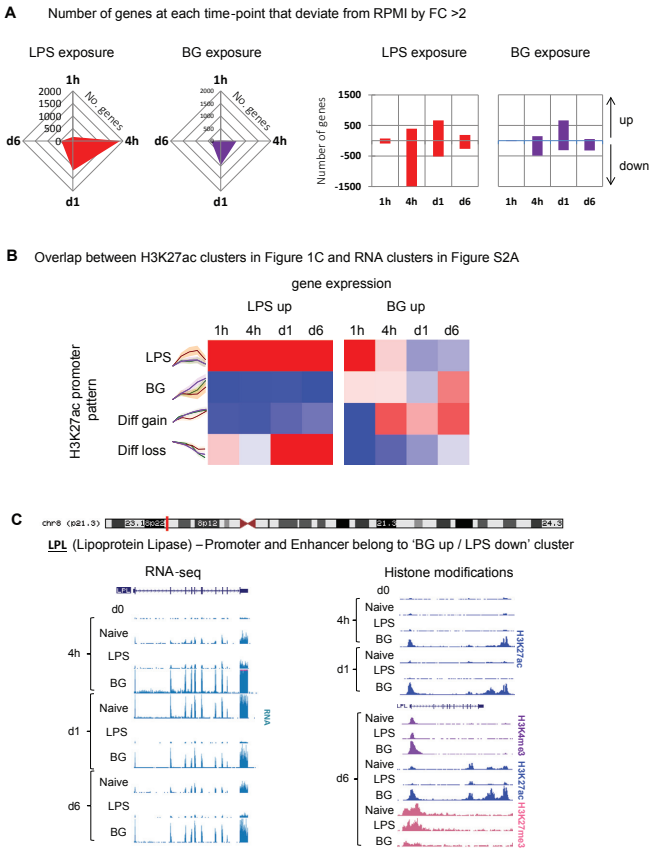


Figure S2. RNA-Seq Dynamics in Response to LPS and BG and Relationship to Histone Marks, Related to Figure 1

(A) Number of genes showing treatment (LPS or BG) specific expression at each time point (1h, 4h, d1, d6). LPS exposure induces the largest number of genes at each time-point, with a minimum of 110 transcripts at 1h, and a maximum of 650 transcripts at day 1. Up to 100 genes maintain LPS-specific expression at d6. Comparatively BG induced gene expression patterns peak at d1, a fraction of which is maintained to d6.

(B) Overlap between gene expression group and promoter H3K27ac cluster. LPS-induced H3K27ac accumulation at promoters correlates well with LPS induced gene expression at all time-points. However, at day 1 and day 6, the 'LPS-up' genes are equally explained by a lag in differentiation-associated repression in LPS treated cells. Conversely, BG exposure leads to faster expression of differentiation associated genes, with higher overlap between 'BG-up' genes and 'differentiation gain' and BG-associated H3K27ac promoters.

(C) Example tracks of a BG induced/LPS repressed gene and an LPS induced gene, LPL (Lipoprotein Lipase).

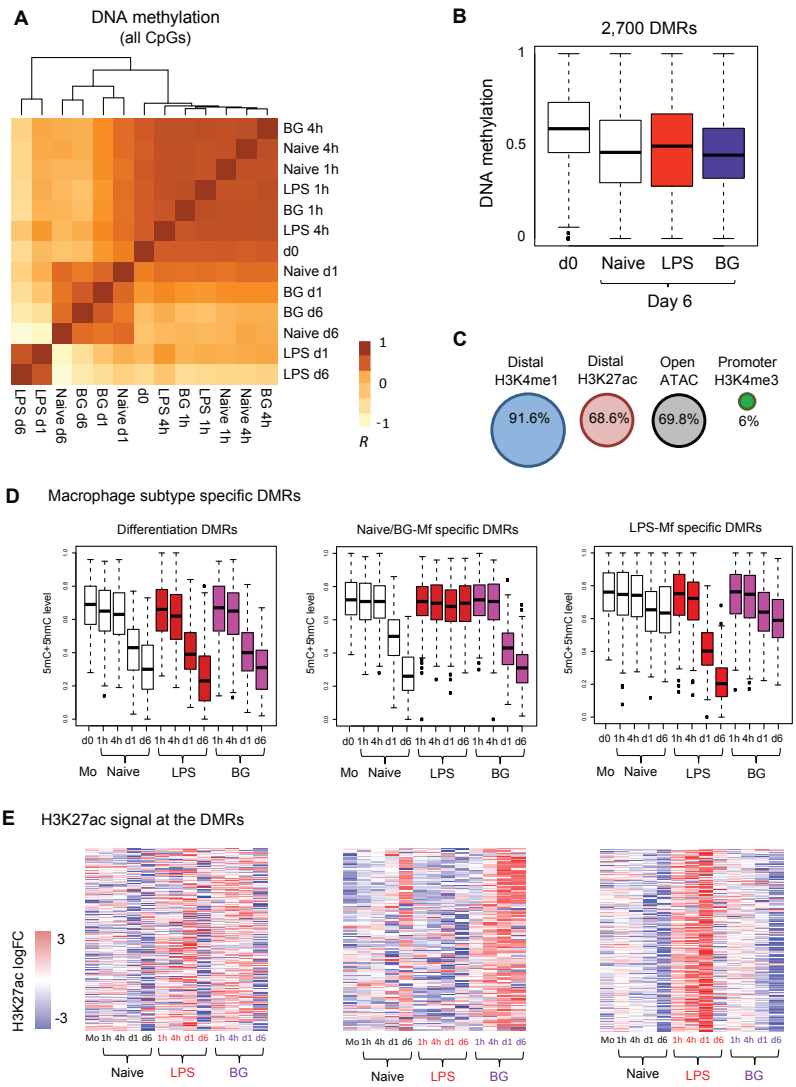


Figure S3. DNA Methylation Dynamics in Monocyte-to-Macrophage Differentiation and Tolerance and Training, Related to Figure 1

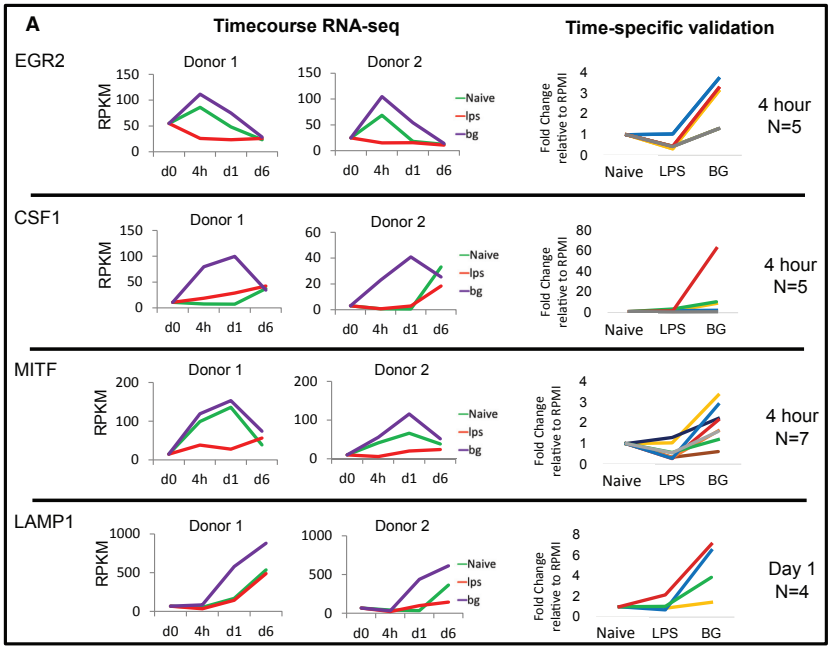
(A) Correlation plot of DNA methylation values, showing clear separation of LPS d1 and LPS-d6 from other samples.

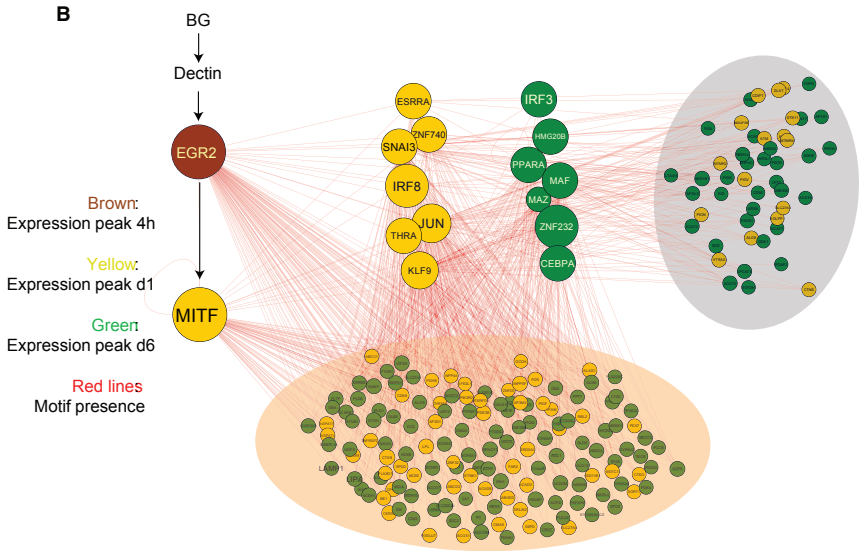
(B) Boxplot of 2,700 DMRs, showing that the general trend is loss of methylation during monocyte-to-macrophage differentiation.

(C) Chromatin context of DMRs. The majority (91%) of DMRs occur in distal regions marked by H3K4me1, 69 occur at H3K27ac marked enhancers and open chromatin regions. Only 6% occur at promoters.

(D) Boxplots showing DNA methylation over time for macrophage sub-type specific DMRs. Analysis identified DMRs common to all macrophages, and those that are only established in LPS-Mf or not-established in LPS-Mf.

(E) Heatmap of H3K27ac changes at DMRs. Generally, DNA de-methylation at DMRs was associated with accumulation of H3K27ac.





(legend on next page)

Figure S4. Expression of Transcription Factors with Enriched Motifs at BG-Associated Promoters and Enhancers and Pathways Associated with Downstream Genes, Related to Figure 2

(A) The expression of main genes enriched at 'BG up / LPS down' and 'Differentiation gain' promoters and enhancers is shown separately for each donor over time. Naive cells are green, LPS exposed cells are red, and BG exposed cells are purple. *EGR2* expression peaks transiently at 4 hr in BG exposed cells, but by day 6, there is no difference between Naive, LPS-Mf or BG-Mf. *CSF1* and *MITF* expression peaks at day 1 and then is reduced. Downstream TF *USF2* shares one motif with *MITF*, and shows high expression in BG macrophages at day 6. *LAMP1* is a major component of the lysosome, and together with *LAMP2* makes up 50% of all lysosomal proteins. *LAMP1* expression peaks late, and is significantly higher in BG-Mf compared to naive and LPS-Mf. qPCR was used to validate RNA-seq results in monocytes from multiple donors.

(B) Transcription Factor network based on EGR2 and MITF motif occurrence at BG induced lysosome and lipid metabolism genes. The size of the nodes represents the number of connections. EGR2 motif is present in the MITF promoters (thick connection). EGR2 and/or MITF motifs are present in another 28 TFs, which themselves have 14 distinct motifs (and are visible as a cluster. Most genes have a combination of EGR2, MITF and a downstream TF motifs (light brown circle). The set of genes to the right do not have EGR2 or MITF motifs, but have motifs for one of the downstream TFs (light gray circle). Overall this network explains 79% of BG-induced lipid metabolism and lysosome-associated genes, compared to 58% based on EGR2 and MITF scan alone. BG induces *EGR2* expression, through its receptor, Dectin-1, and higher expression of *MITF* is observed, as well as its activator cytokine factor *CSF1* (see also Figure S4). Conversely, LPS treatment represses *EGR2*, *CSF1* and *MITF*. Genes are labeled by time at which their expression peaks in BG exposed cells. *EGR2* expression peaks at 4 hr (brown), *MITF* and *KLF9* at day 1 (gold). The rest of the downstream genes peak at day 1 (gold) or peak at day 6 (green). Connections between TFs and downstream genes is shown as red lines.

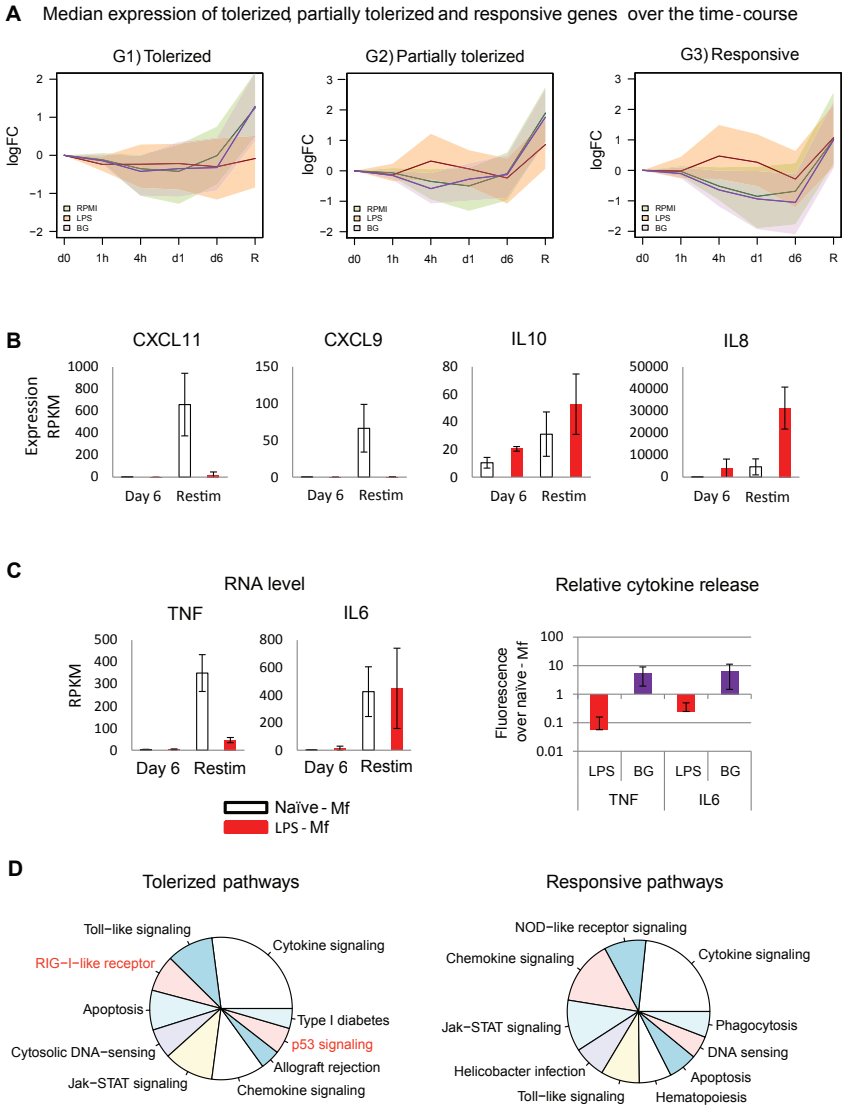


Figure S5. Tolerance at the Transcriptional Level, Related to Figure 3

(A) Pattern of expression of tolerized and responsive genes during the time-course shown as median logFC of two donors (with first and third quartiles shown as shaded areas). The most tolerized (G1) genes did not show upregulation in response to the initial LPS exposure in monocytes, while responsive genes (G3) showed high induction in monocytes.

(B) Notable examples of tolerized and responsive genes. Data are shown as mean RPKM and error bars are standard deviations. Data are represented as mean ± SD.

(legend continued on next page)

(C) Expression of *IL6* and *TNF*. Release of these proteins from macrophages in response to LPS is considered the gold-standard for determining tolerance. At the transcriptional level TNF is partially tolerized, while IL6 is responsive in LPS-Mf. Error bars represent standard deviation. IL6 and TNF protein release after LPS restimulation is high in BG-Mf and absent in LPS-Mf compared to naive-Mf. The disconnect between transcription and release of IL6 can potentially be explained by the larger size and higher lysosome content in BG-Mf, induced by early activation of lipid and lysosome pathways in BG exposed cells.
(D) Top 10 KEGG pathways enriched in tolerized and responsive gene groups from DAVID ontology analysis. Area relates to the number of genes within the pathway, red font signifies that the pathway only shows significant enrichment in the tolerized gene group. Cytokine-cytokine receptor signaling was the top pathway in both tolerized and responsive groups indicating that cytokine genes are equally spread across the gradient of LPS-Mf response to LPS re-exposure.

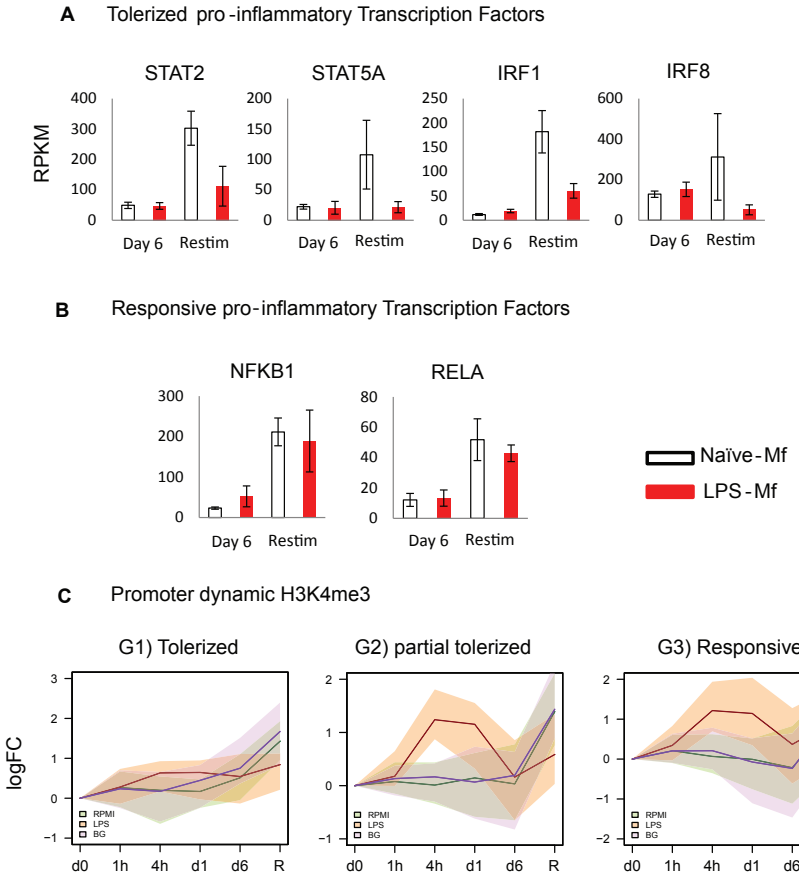


Figure S6. Active Histone Mark Changes at Promoters of Tolerized and Responsive Genes and Overall Chromatin States at the Same Promoters, Related to Figure 4

(A) Expression at day 6 and at LPS re-exposure for STAT2 and −5A, and IRF1 and −8 (mean RPKM of 4 donors, error bars represent standard deviation). These pro-inflammatory TFs show a tolerized response in LPS-Mf to LPS re-exposure. The inability of these genes to be activated may play a role in the tolerance of downstream targets, as suggested from the enrichment of their motifs in the G2 partially tolerized gene promoters (Figure 4B).

(B) expression at day 6 and at LPS re-exposure for NFKB1 and RELA. These TFs are responsive to LPS re-exposure in LPS-Mf, and their motifs are not significantly enriched in tolerized genes. This suggests that NF-κB signaling is not impaired at the level of transcription. Data are represented as mean ± SD.

(C) LPS-Mf do not accumulate H3K4me3 at tolerized genes, but do so at the promoters of responsive genes. This pattern is similar to that of H3K27ac shown in Figure 4D.

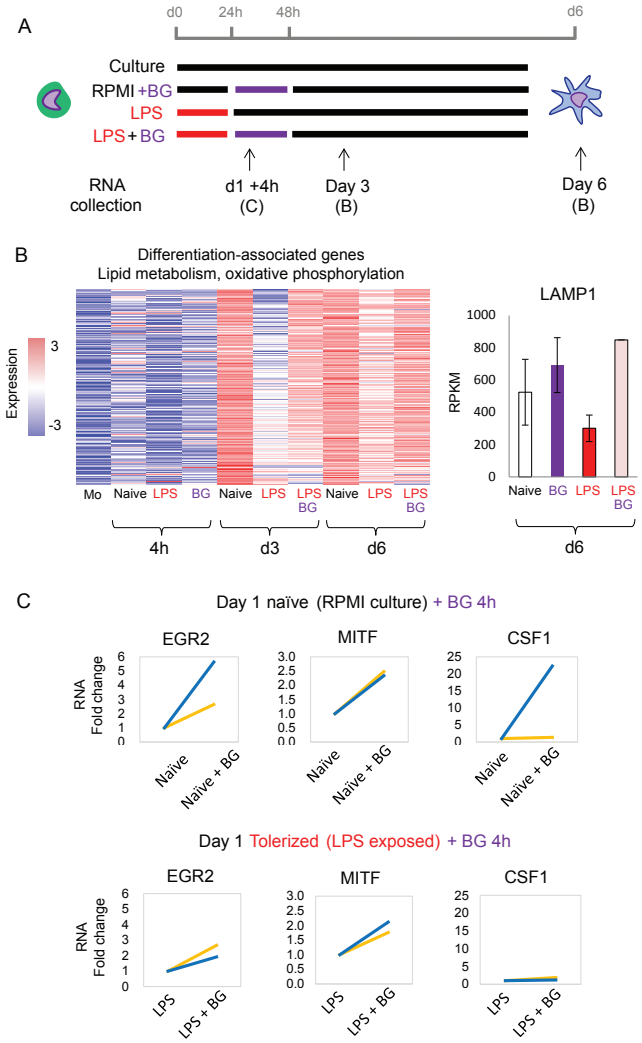


Figure S7. Expression of Genes Involved in Lipid Biosynthesis and Metabolism following BG Reversal of LPS-Induced Tolerance, Related to Figure 7

(A) Experimental set-up, indicating the collection of samples for gene expression analysis. Samples were collected at day 1 +4h, indicating that monocytes were treated with media (RPMI) or LPS for 24 hr, at which point cells were exposed to BG for 4 hr and collected. Additionally samples were collected at day 3 and day6.

(B) BG exposure, following LPS, recovers the expression of genes involved in lipid biosynthesis and oxidative phosphorylation as early as day 3. LAMP1 is an example of a lysosome gene that shows high expression in BG-Mf and low expression in LPS-Mf. BG exposure recovers the expression of this gene in LPS-BG-Mf.

(C) BG addition at day 1 in Naive monocytes induces the expression of EGR2, MITF and CSF1, as it does when added at day 0 (Figure S4C). In tolerized monocytes, BG induces the expression of EGR2 and MITF, but to a lesser degree. This indicates that BG receptor pathways are not completely disrupted by LPS exposure, providing a basis for BG reversal of LPS-induced tolerance.

[1]Radboud University, Department of Molecular Biology, Faculty of Science Nijmegen, the Netherlands.
[2]Roche NimbleGen, Incorporated, 500 South Rosa Road, Madison, WI 53719, USA
[3]Nuclear Dynamics Programme, The Babraham Institute, Babraham Research Campus, Cambridge, CB223AT, UK.
[4]These authors contributed equally to this work.
[5]Corresponding author

Chapter 5

Higher-order chromatin reorganization during monocyte-to-macrophage differentiation

Shuang-Yin Wang[1,4], Tatyana Kuznetsova[1,4], Boris Novakovic[1], Koen H.M. Prange[1], Onkar Joshi1, Cheng Wang[1], Robab Davar[1], Eva M. Janssen-Megens[1], Todd Richmond[2], Daniel Burgess[2], Biola Javierre[3], Mikhail Spivakov[3], Peter Fraser[3], Hendrik G. Stunnenberg[1,5]

ABSTRACT

Monocyte-to-macrophage differentiation *in vivo* occurs in homeostasis and during inflammation. This differentiation is associated with changes on epigenetic and transcriptional levels. Here we investigate dynamics of higher-order chromatin organization during monocyte-to-macrophage differentiation and immunological training by Capture Hi-C. We demonstrate that despite the absence of cell division, this differentiation is associated with higher-order chromatin reorganization, including changes in A/B compartments, strengthening of TADs and their boundaries and a global gain of short-range interactions. While differentiation is associated with profound changes, immunological training does not cause spatial chromatin dynamics, suggesting that differentiation is the main driver of chromatin remodeling.

INTRODUCTION

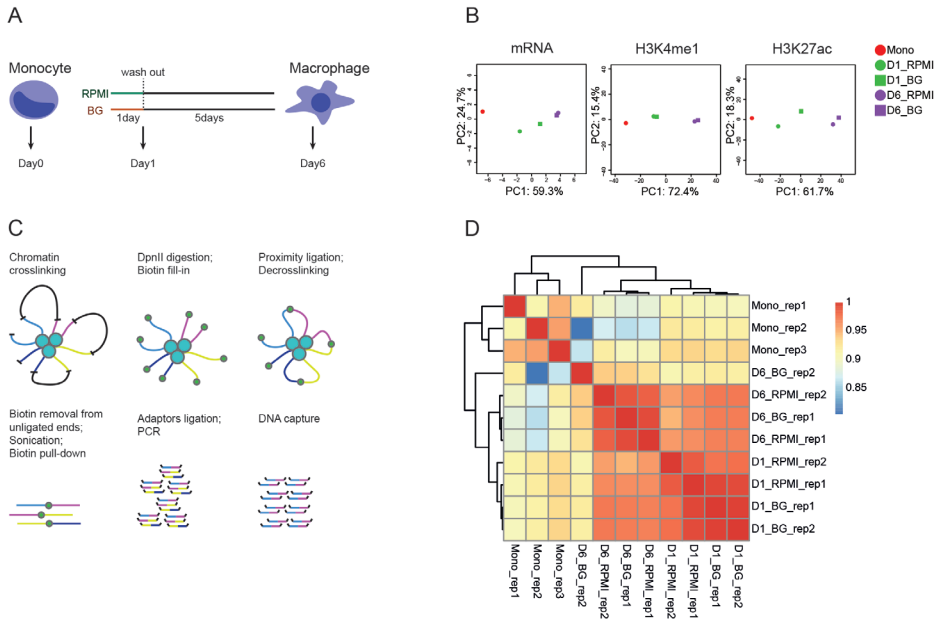
Under homeostatic conditions monocytes can differentiate to macrophages in vitro and in vivo (1). In steady state condition the contribution of circulating monocytes to tissue macrophages is limited. However, upon inflammation, monocytes are rapidly recruited and give rise to tissue macrophages, which can effectively combat an infection. Upon encountering different inflammatory stimuli these macrophages display stimulus-specific responses, suggesting their functional plasticity(1-3). Furthermore, it has been recently discovered that in response to certain stimuli (e.g. *Candida Albicans* β-glucan) monocytes can develop an enhanced immunological response and resistance to secondary infections – a phenomenon described as "trained innate immunity" (4,5).

Monocyte-to-macrophage differentiation and response to inflammatory stimuli are associated with extensive remodeling of the gene expression program(2,5,6). These changes are largely driven by combinatorial action of lineage-specific (e.g. *PU1, EGR2*) and signal-dependent transcription factors (e.g. *NF-ĸB, IRF*) that bind and activate distal regulatory elements or enhancers(7,8). Accordingly, the enhancer repertoire is cell-type specific, and in macrophages was shown to be shaped by the tissue microenvironment and exposure to external stimuli(3,9,10).

Regulatory activity of enhancers can be exerted over large linear distances and the three-dimensional organization of chromatin controls proper communication between regulatory elements. Recently, it has been demonstrated that on the Megabase (Mb) scale chromatin is organized into two compartments: the 'A' compartment that is generally associated with an active chromatin configuration and the 'B' compartment containing repressive chromatin(11,12). On a finer scale the A and B compartments are further partitioned into self-interacting units or topologically associating domains (TADs). The majority of long-range interactions, including those between regulatory elements, occur within TADs while interactions with other TADs are relatively infrequent(13,14).

Here, we set out to characterize chromatin organization during terminal myeloid differentiation and immunological training by focusing on two closely related cell types: monocytes and macrophages. We report widespread changes in 3D chromatin organization during monocyte-to-macrophage differentiation. With the use of the high-resolution capture Hi-C (CHi-C) approach(15,16), we show that monocyte-to-macrophage differentiation is accompanied by extensive 'A/B' compartment switching, strengthening of TAD boundaries and gain of short-range interactions between regulatory elements, while immunological training does not involve changes in higher-order 3D chromatin organization.

Figure 1.



RESULTS

Capture Hi-C approach

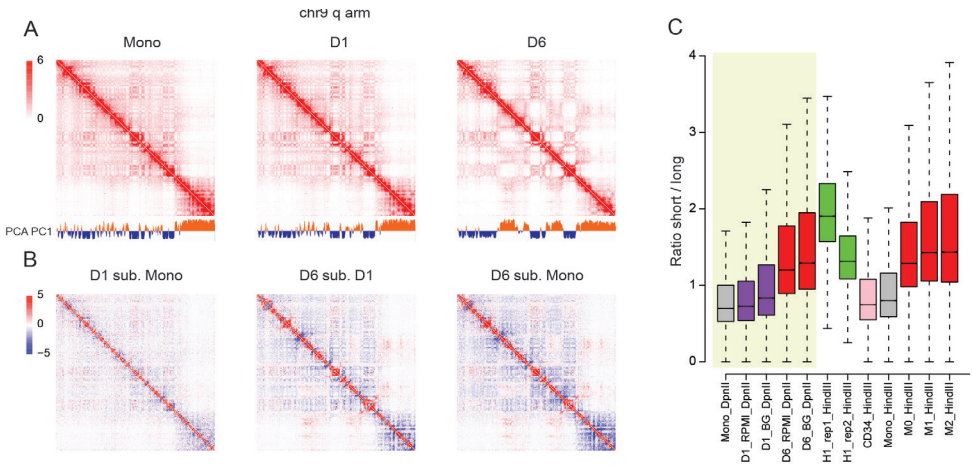
To study how monocyte-to-macrophage differentiation and immunological training affect chromatin organization, human primary monocytes were isolated from healthy volunteers and differentiated into naïve macrophages by ex-vivo culturing for 6 days (5). Immunological training was performed by an initial exposure of the monocytes to β-glucan (BG) during the first 24 hours, followed by withdrawal of the stimulus and 5 days of culture (Figure 1A). We have previously demonstrated that monocyte-to-macrophage differentiation is associated with extensive epigenetic and transcriptional remodeling, with a distinct signature associated with different time-points during differentiation (Figure 1B) (Novakovic, et al). In order to investigate chromatin organization during differentiation and immunological training we performed CHi-C on monocytes (Day0), intermediate population (Day1/D1) and differentiated macrophages (Day6/D6), using the four basepairs (bp) cutting restriction enzyme DpnII. CHi-C combines genome-wide detection of chromatin interactions with target enrichment by hybridization-based capture, allowing for high-resolution detection of interactions at a selection of loci (Figure 1C). Target selection was based primarily on H3K27Ac dynamics

during monocyte-to-macrophage differentiation, as this mark is associated with active regulatory elements. We have targeted 20193 enhancers and 6826 promoters, ~40% and ~50% of which, respectively, were associated with a dynamic H3K27ac signal.

All CHi-C libraries were deep-sequenced to obtain at least 250 million reads per library, which resulted in at least 100 million informative reads after filtering (Supplementary Table1). The enrichment of the targeted regions was ~52 fold and average capturing efficiency was 56%. The uncaptured remainder of paired-end reads (44%) in our CHi-C samples represented conventional Hi-C reads.

To validate the robustness of the CHi-C approach, libraries for each time-point were prepared in two biological replicates, which displayed high reproducibility (data not shown). Furthermore, global clustering of interactions revealed high similarity between samples belonging to the same time-points (Figure 1D).

Figure 2.



Global reorganization of 3D chromatin architecture during monocyte-to-macrophage differentiation

To investigate the higher-order chromatin structure globally, we analyzed our CHi-C data as conventional Hi-C and compared the interaction frequencies on the chromosomal level. Interestingly, in monocytes (Figure 2A, left) and Day1 cells (Figure 2A, middle), we observed frequent dispersed long-range interactions, resulting in a blurred pattern on the 2D interaction matrix. In Day6 macrophages, these long-range interactions were largely lost. The retained interactions localized to specific regions, resulting in a sparse checkerboard pattern (Figure 2A, right). To further investigate the observed contact redistribution we

directly compared the differentiation time-points by subtracting Hi-C interaction frequencies (Figure 2B). We observed a progressive increase in short-range (diagonal of the interaction matrix) and a clear decrease in long-range interactions during differentiation (Figure 2B).

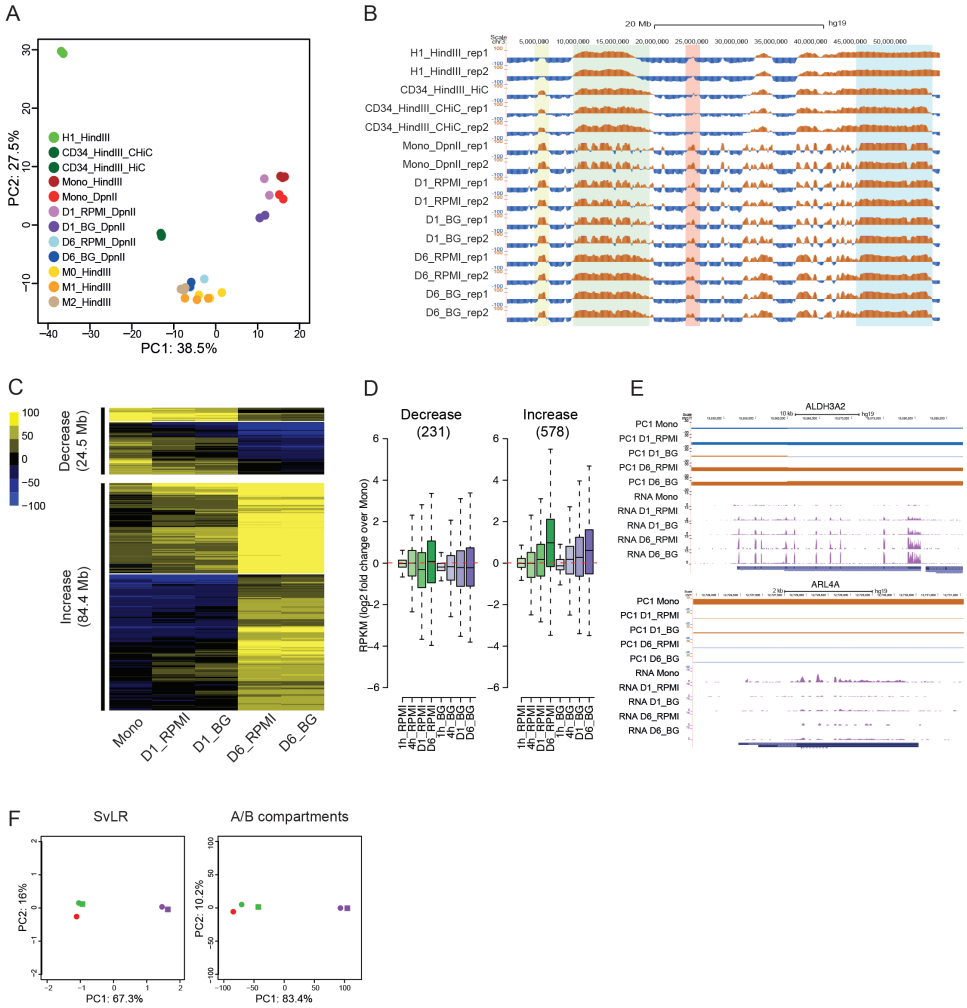
To quantify the observed redistribution of chromatin contacts we calculated the ratio of short (<2Mb) versus long-range (>2Mb) interactions (SvL)(17). As expected, Day6 macrophages displayed higher SvL (1.91 fold) than monocytes and Day1 cells (Figure 2C, yellow box). Next we were wondering whether the increase in SvL ratio is associated with differentiation in general or is specific for monocyte-to-macrophage differentiation. In order to address this question we analyzed external Hi-C and CHi-C datasets for several human cell types: human embryonic stem cell (hESC) line H1 (12), hematopoietic progenitors CD34+(15), monocytes and three subtypes of macrophages (M0, M1 and M2) (Javierre et al., Cell in press). Interestingly, comparison of these cell types (H1, CD34, Monocytes) revealed a general decrease in SvL ratio during differentiation, in contrast to an increase observed during monocyte-to-macrophage differentiation. Furthermore, the high SvL detected in naïve macrophages at Day6, is consistent across different macrophage subtypes (M0, M1, M2; data from Javierre et al., Cell in press) (Figure 2C). This indicates that a high SvL ratio, as compared to monocytes, is a general property of macrophages.

A/B compartment rearrangement in mature macrophages

On a megabase (Mb) scale the genome can be partitioned into two compartments, A and B, which contain relatively active and inactive chromatin, respectively, and preferentially interact with regions belonging to the same compartment. Genes residing in the 'A' compartment tend to be expressed at higher levels than those in the 'B' compartment(11). 'A/B' compartment organization is cell-type specific and ~36% of the genome was shown to change compartments in at least one lineage during hESC in vitro differentiation(12).

To globally assess chromatin organization dynamics during myeloid differentiation we performed principal component analysis (PCA). As expected, hESC clustered away from CD34 and differentiated myeloid cells, indicating that differentiation is accompanied by reorganization of chromatin. Within the differentiated cells we detected two distinct groups, separated on both principal components. The first group included monocytes together with Day1 cells. The second group contained fully differentiated macrophages. All macrophage subtypes displayed tight clustering, irrespective of their subtype, indicating the similarities in their chromatin organization (Figure 3A).

Figure 3.



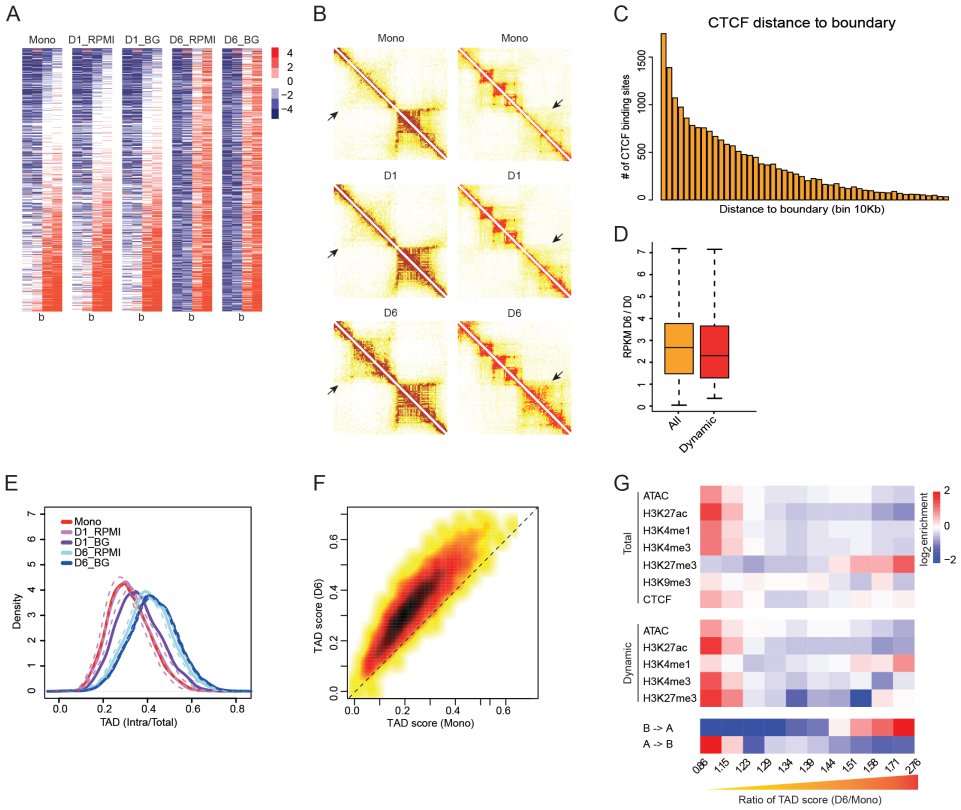
Next, we identified 'A/B' compartments based on the first principal component (PC1) from the PCA; positive PC1 values were assigned to the 'A' compartment (orange) and negative values to the 'B' compartment (blue) (Figure 3B). Visual inspection confirmed cell-type specific changes in compartment organization during myeloid differentiation. For example, we observed regions changing compartments early, during H1 to CD34 differentiation and remaining in the same compartment in differentiated myeloid cells (Figure 3B, yellow box), those changing compartment in monocytes and remaining in the same compartment in macrophages (Figure 3B, red box), or unchanged throughout differentiation (Figure 3B, blue box).

Several studies suggest that chromatin organization is constrained during interphase(18,19) and mitosis is required for a region to be repositioned(20). Remarkably, while monocyte-to-macrophage differentiation occurs without cell division(5), we observed multiple regions switching compartments (Figure 3B, green box). In total ~3,3% of the genome (108,9 Mb) was associated with significant changes in PC1 values during monocyte-to-macrophage differentiation, with the majority of the regions (~2,6%, 84,4 Mb) relocating from 'B' to 'A' or showing a strong increase in PC1 values within the 'A' compartment; and ~0,7% (24,5 Mb) relocating from 'A' to 'B' compartment or associated with decrease in PC1 value within the B compartment (Figure 3C). Interestingly, while monocyte-to-macrophage differentiation is associated with prominent 'A/B' compartment reorganization, treatment with β-glucan has marginal effect (Figure 3A-C). We could identify only two regions specific for β-glucan treated macrophages at Day1 and none at Day6 (Supplementary Figure 1).

As the 'A/B' compartments are suggested to be associated with relative activity state, we next analyzed the expression of genes located in changing compartments. As expected, genes residing in the regions with decreasing PC1 values were associated with reduction in gene expression during differentiation (Figure 3D, left), while those associated with increased PC1 values showed gradually induced expression levels (Figure 3D, right). For example, expression of the *ALDH3A2* gene is induced more than 15 fold during differentiation; accordingly, we observe this gene relocating from the inactive 'B' compartment in monocytes to the active 'A' compartment in macrophages (Figure 3E, top). On the other hand, expression of the *ARL4A* gene decreases 2 fold during differentiation, and the gene is relocated to the inactive B compartment (Figure 3E, bottom). Global gene ontology (GO) analysis of genes residing in switching compartments did, however, not reveal any enrichment in function for these genes, suggesting that multiple pathways might be affected by compartment reorganization.

Lastly, we were wondering when the observed changes in higher order chromatin organization happen with respect to the epigenetic and transcriptional changes during the differentiation process that we previously identified (Figure 1B). For this, we performed PCA analysis of SvL ratio and 'A/B' compartment organization for monocytes, Day1 and Day6 naïve and β-glucan trained macrophages. We observed that while transcriptional and epigenetic changes were already prominent at Day1 (Figure 1B), changes in higher-order chromatin organization were only detected at Day6 (Figure 3F), suggesting the temporal uncoupling of these events. Furthermore, while on the level of gene expression and H3K27ac occupancy we detected a clear difference between naïve and β-glucan exposed macrophages at Day1 (Figure 1B), we did not detect treatment-dependent changes in higher order chromatin structure (Figure 3F).

Figure 4.



Strengthening of TAD boundaries during monocyte-to-macrophage differentiation

On a Kilobase scale the 'A/B' compartments are further partitioned into TADs. Multiple studies suggest that TAD organization is similar between cell types and stable during differentiation (13,14), while others demonstrate that TAD boundaries can be altered (21,22). In order to investigate the TAD organization during monocyte-to-macrophage differentiation we first performed TAD calling based on the 40 Kb resolution interaction matrix. We detected ~2000 TADs in monocytes and ~3000 TADs in macrophages. The large difference in the number of TADs that we identified in monocytes and macrophages suggests extensive remodeling of TAD boundaries. To investigate this, we assessed the boundary strength by computing and comparing the directionality index (DI) of monocytes and macrophages (13). 187 boundary regions were identified as dynamic based on 4 fold change cut-off of DI. In agreement with the increased number of TADs in macrophages, all of the dynamic boundaries were stronger at Day6 (Figure 4A-B). Furthermore, increase in boundary strength was detected almost

exclusively at Day6 and not at Day1, consistent with the 'A/B' compartment reorganization (Figure 3C).

As TAD boundaries are frequently associated with insulator protein CTCF (13) and CTCF is required for their maintenance (23), we hypothesized that strengthening of the TAD boundaries during monocyte-to-macrophage differentiation is associated with differential CTCF binding. Motif analysis of TAD boundaries confirmed the frequent association of the CTCF motif with TAD boundaries in monocytes and macrophages (data not shown). Thus, we performed CTCF ChIP-seq in monocytes and macrophages and compared the binding profiles. As expected, CTCF binding was enriched at TAD boundaries (Figure 4C), however, no significant difference in CTCF occupancy was found at macrophage-specific boundaries (Figure 4D). Hence, the increase in boundary strength in macrophages is not caused by increased CTCF occupancy in macrophages. Motif enrichment analysis at dynamic boundaries did not reveal any additional motif besides CTCF, associated with these boundaries.

Terminal myeloid differentiation alters intra-TAD chromatin organization

The higher number of called TADs in macrophages might also suggest changes in chromatin organization within the TADs. In order to investigate this, we computed the TAD score – a ratio of intra-TAD interactions to inter-TAD interactions on the same chromosome (12,24). For most of the TADs we observed a gradual increase in TAD score during monocyte-to-macrophage differentiation (Figure 4E-F), suggesting that domains become more compartmentalized in macrophages. Next, we wanted to assess how the increase in TAD score relates to changes in the epigenetic and transcriptional landscape. For this, we have sorted the TADs based on the fold change of their score in macrophages compared to monocytes and divided them in ten bins. Bin one contained the least and bin ten the most changing TADs. These dynamic TADs were not enriched for differentially expressed genes (data not shown), suggesting that increased compartmentalization of these TADs is not directly associated with transcriptional changes of genes within them.

We reasoned that while we could not detect a distinct gene expression pattern associated with dynamic TADs, the changes in TAD score might be associated with epigenetic changes. For each bin we calculated the enrichment of activating (H3K27Ac, H3K4me1, H3K4me3) and repressive (H3K27me3, H3K9me3) histone marks, accessible chromatin and insulator protein CTCF (Figure 4G). Interestingly, the least changing TADs appeared to be associated with activating marks, open chromatin and switching from active to inactive compartment. The most changing TADs on the other hand, were found to be enriched for the repressive histone mark H3K27me3 and switching from inactive to active compartment (Figure 4G). One possible scenario is that upon relocation from inactive to active compartment, these TADs are restricted to prevent the spread of H3K27me3 or long-range repression of regulatory elements in neighboring TADs.

DISCUSSION

In this study we analyzed changes in chromatin organization during terminal myeloid differentiation and induction of trained immunity. Notably, monocyte-to-macrophage differentiation occurs in the absence of cell division (5), allowing us to uncouple differentiation-induced changes in chromatin organization from changes associated with cell division.

During differentiation changes in gene expression are accompanied by reorganization of chromatin (12,25). As these changes often coincide, it is not clear whether changes in gene expression and nuclear repositioning are causally related. Some locus-specific studies suggested that changes in gene expression and epigenetic make-up are sufficient to induce repositioning of the locus (26,27), while others suggest that changes in gene expression do not lead to repositioning (28). During monocyte-to-macrophage differentiation extensive epigenetic and transcriptional changes can already be observed at Day1 (Figure 1B). Remarkably, no changes in higher-order chromatin organization could be detected at this time-point (Figure 3F). This temporal uncoupling of chromatin reorganization and transcriptional/epigenetic changes might suggest their independence and lack of a direct causal relation, although further experimental validation is required to confirm this observation.

Changes in 'A/B' compartments during differentiation have been recently reported for hESC and several derived lineages (12). The extent of the observed changes varied dependent on the assayed lineage, from ~3,8% in mesendoderm to 25% in mesenchymal stem cells (MSC)(12). In monocyte-to-macrophage differentiation, despite of the close relation between the two assayed cell types, ~3,3% of the genome was associated with compartment changes, comparable with the extent of changes between hESC and mesendoderm. Interestingly, during hESC differentiation the majority of the detected compartment changes were associated with an 'A' to 'B' transition and expansion of the inactive 'B' compartment(12), while in monocyte-to-macrophage differentiation the majority of the switching regions relocated from the 'B' to the 'A' compartment. Recent reports suggest that stimulus-specific responses during macrophage activation are associated with distinct transcription factor networks, which are signal-dependent and operate in a permissive chromatin environment (2,27,29). Thus, the general increase in PC1 values and switching to the 'A' compartment can be associated with the establishment of an open, permissive transcriptional environment in macrophages, which might attribute to their functional plasticity (3). In line with this, our analysis revealed that macrophage subtypes, while different on transcriptional, epigenetic and functional levels (2,29), share a similar chromatin organization (Figure 3A).

Coinciding with a global gain of short-range interactions and switching to the 'A' compartment, we observed that topological domains in macrophages adopt a more pronounced organization (Figure 4A-B,E-F). This reorganization is accompanied by strengthening of TAD boundaries, which appears to be CTCF-independent (Figure 4A-D). However, other insulator proteins might be responsible for increase in boundary strength.

For instance, a component of the cohesin complex, SMC1A, is strongly up-regulated (~3 fold) during monocyte-to-macrophage differentiation (data not shown). The cohesin complex is required for insulator activity of CTCF(30) and its depletion leads to reduction of intra-TAD interactions(31), in agreement with pattern observed in monocytes, where SMC1A is lowly expressed. Alternatively, boundary strength might be affected by combinatorial action of insulator proteins (32), similar to boundary remodeling in response to temperature stress in Drosophila (21).

Functionally, the most changing TADs display strong enrichment of the H3K27me3 chromatin mark (Figure 4G). We speculate that stronger insulation of these TADs might be required to prevent the spreading of repressive chromatin or, alternatively, to prevent an aberrant activation of inactive genes within the domain, as has been recently reported for the HOX locus (23). Future genome editing experiments of dynamic TAD boundaries are required to address their functional significance in macrophage biology.

In summary, we show that monocyte-to-macrophage differentiation is accompanied by extensive changes in three-dimensional chromatin organization, including a global increase in short-range interactions in macrophages, changes in 'A/B' compartments and reorganization of TADs. Treatment with β-glucan, on the other hand, does not alter higher-order chromatin organization, suggesting that differentiation, rather than treatment-specific activation is the main driver of changes in the three-dimensional chromatin structure of macrophages.

Authors' contributions

TK performed the experiments, prepared the figures and wrote the manuscript; SYW analyzed the data, prepared the figures and edited the manuscript; BN performed ChIP-seq and RNA-seq experiments; KP contributed to the analysis and edited the manuscript; OJ, CW and RD contributed to the experiments; EMJM carried out the sequencing of the libraries; TR and DB provided capture libraries; BJ, MS and PF assisted with cHiC experiments and data analysis. HGS designed and supervised the project and wrote the manuscript.

MATERIALS AND METHODS

Monocytes isolation from healthy donors

All primary cells were isolated from healthy volunteers who gave written informed consent (Sanquin Blood bank, Nijmegen, the Netherlands). Peripheral blood mononuclear cells were isolated by centrifugation in Ficoll-Paque (GE Healthcare), followed by removal of T cells using an additional Percoll gradient. Monocytes were purified from PBMCs using negative selection in an LD column magnet separator, with beads for CD3+ (T cells), CD19+ (B cells) and CD56+

(NK cells) positive cells (Miltenyi biotech), yielding >95% pure monocytes. Successful isolation of monocytes was confirmed with FACS, as previously described (5).

In vitro Monocyte to macrophage differentiation and induction of trained innate imunity

Monocytes were differentiated into resting macrophages by ex vivo culture in RPMI 1640 medium (Sigma Aldrich) with 10% Human Serum. Trained innate immunity was induced by treatment with 5 µg/mL BG for 24 hours, followed by washout and 5 days in culture. Establishment of tolerance or training in the resulting macrophages at day 6 was determined by TNF and IL6 release at 24 hours following LPS stimulation using ELISA. For ChIP-seq, 10x106 monocytes were seeded in 10cm dishes, for CHi-C at least 50x106 monocytes were seeded in 14cm dishes.

Chromatin immunoprecipitation and sequencing (ChIP-seq)

ChIP was performed according to standard protocol (33) with minor modifications. Paraformaldehyde (1%) crosslinking was carried out for 10 minutes followed by the chromatin preparation as described earlier (34). Nuclei were re-suspended in ChIP-incubation buffer at a concentration of $20x10^6$ cells/ml and sheared (5 cycles with each cycle containing 10 sec power on and 10 sec interval) using BioruptorPico (B01060001, Diagenode). Sonicated chromatin equivalent of $4x10^6$ cells was incubated with 15 µl CTCF antibody (ab70303, Abcam) overnight at 4°C. ChIP-Seq sample preparation and sequencing was performed according to manufacturer's instructions (Illumina).

Capture Hi-C (CHi-C)

The Capture Hi-C experiment was divided into two parts, in-nucleus Hi-C and ss-DNA probe capture enrichment. In-nucleus Hi-C was carried out as described in Nagano et al., 2015. DpnII was used as the restriction enzyme. On beads DNA amplification PCR was carried out with 7-9 cycles to generate around 1µg of Hi-C library DNA. ssDNA probe capture step was carried out using the protocol provided by Roche NimbleGen Inc. (http://sequencing.roche.com/products/nimblegen-seqcap-target-enrichment/seqcap-ez-system/seqcap-ez-developer.html) optimized for the probe capture library. Libraries were indexed using NEXTflex adapters (Bioo-Scientific Corporation, Austin, TX, USA) and 75bp or 43bp paired-end sequencing was performed on Illumina instruments using TruSeq reagents Illumina, San Diego, CA, USA), according to manufacturer's instructions.

ChIP-Seq data analysis

Tags were mapped to the reference human genome hg19 using the Burrows-Wheeler Alignment Tool[2] (BWA). Before down stream use, duplicates reads and reads with a MAPQ<15 were discarded. For visualization the number of overlapping sequence reads was determined per base pair, averaged over a 10 bp window and visualized in the UCSC genome browser (http://genome.ucsc.edu). For comparison bam files were normalized to the same number of reads. Peakcalling was performed with MACS2((35)) software and peaks for monocytes and macrophages were merged. RPKM was calculated for merged peaks.

Capture Hi-C data analysis

Mapping

The two ends of paired-end reads were mapped to the reference human genome (hg19) separately, using BWA MEM with default parameters. Reads were filtered based on mapping quality score (both ends MAPQ≥10) and PCR duplicates were removed. Reads were also removed if the two ends are from the same DpnII fragment.

Normalization

First, all the reads were treated as from conventional Hi-C library. Raw contact matrices were generated for individual libraries and merged libraries from the same condition at fixed locus sizes of varying resolution including 20kb, 40kb, 100kb, 200kb. For each matrix, the contacts on the diagonal bins were considered informative and removed. Normalization of the matrices were then performed with iterative correction (36) in Bioconductor R package “HiTC” with default settings. After normalization, the rows of the contact matrix approached a constant. Differential Hi-C interaction maps were calculated by subtracting the second condition from the first one using the normalized contact matrices of 200kb resolution.

SvL ratio calculation

The short-range versus long-range, or SVL ratio, was calculated as described previously(17). In this calculation, we considered only intra-chromosomal interactions within a single chromosome using the normalized contact matrices of 40kb resolution.

A/B compartment identification

A/B compartment calling was performed at 40kb resolution as previously described (12). Briefly, the normalized 40kb contact matrices were used to calculate the expected contact frequency

between two 40-kb bins given the distance separating them in the genome. A sliding 200kb window was then used with a step size of 40kb to generate an observed/expected matrix. This matrix was then used to calculate the Pearson correlation matrix and subsequently used for principal component analysis. Genomic regions were identified displaying significant changes in A/B compartment, if these regions show statistically significant variability in PC1 values across all conditions using ANOVA and at least 40 difference in the PC1 values in any two conditions.

TAD calling

Topological domains were systemically identified based on the directionality index (DI) score using a Hidden Markov Model (HMM) previously described (13)at 40kb resolution. The final coordinates of topological domains were manually inspected and calibrated according to all the domains obtained above. The intra-domain contact score (TAD score) was calculated as described previously (25).

Capture Hi-C

CHiCAGO CHi-C analysis package (http://regulatorygenomicsgroup.org/chicago) was used to call significant contacts. To improve quality, the whole genome was windowed into 3 DpnII site tiles. Virtual baits were created using the 3 DpnII site tiles that contain at least one targeted DpnII fragment. Interactions were filtered based on interaction score and number of reads (score≥5 and at least 2 reads).

FIGURE LEGENDS

Figure 1. Three-dimensional chromatin organization profiling during monocyte-to-macrophage differentiation. (A) Schematic overview of experimental set-up for studying monocyte differentiation and trained immuntity. (B) PCA plots of mRNA, H3K27ac and H3K4me1 dynamics during monocyte-to-macrophage differentiation. (C) Schematic representation of CHi-C approach. (D) Global clustering of Hi-C interaction profiles.

Figure 2. Global reorganization of interaction landscape. (A) Normalized Hi-C interaction maps for the q-arm of chromosome 9 in monocytes (left), Day 1 (middle), and Day 6 (right). The color maps for relative interaction probability are displayed on the same scale for each heatmap. The PC1 signal was used to define the A (orange) and B (blue) compartments and is displayed below each heatmap. (B) Differential heatmaps for the q arm of chromosome 9 in monocytes (left), Day 1 (middle), and Day 6 (right). The color maps are displayed on the same scale for each comparison. Red is used to designate enrichment in the first condition, and blue depletion. (C) The ratio of short-range (<2 Mb) versus long-range (>2Mb) interactions (SvL) calculated for each chromosome arm across the whole genome at 40kb resolution.

Figure 3. A/B compartment switching during differentiation. (A) Comparative PCA analysis of chromatin organization during myeloid differentiation. (B) Screenshot exemplifying cell-type specific A/B compartment organization. Region switching in H1 to CD34+ transition (yellow), monocyte-to-macrophage transition (green), CD34+ to monocyte transition (red) and non-dynamic (blue). (C) Heatmap of regions significantly changing their PC1 value during monocyte-to-macrophage differentiation. “Decrease” corresponds to regions with a decrease in PC1 in macrophages compared to monocytes. These regions relocate from A to B compartment or display a significant decrease in PC1 within the B compartment. “Increase” corresponds to regions with an increase in PC1 in macrophages compared to monocytes. These regions relocate from B to A compartment or display a significant increase in PC1 within the A compartment. (D) Box plots of average expression of genes residing in regions with decreased (left) or increased (right) PC1 value. (E) Example screen shot of genes switching compartments during monocyte-to-macrophage differentiation. ALDH3A2 gene (left) is belongs to B compartment in monocytes and Day1 macrophages, and relocates to A compartment in Day6 macrophages. ARL4A gene (right) belongs to A compartment in monocytes and Day1 cells and repositions to B compartment in Day6 macrophages. (F) PCA analysis of time-dependent changes in higher-order chromatin organization. SvL ratio (left) and A/B compatments (right) demonstrate that Monocytes and Day1 macrophages are very similar, while Day6 macrophages are different.

Figure 4. Domain-level changes in chromatin organization. (A) Heatmap displays directionality index (DI) values at 40 Kb resolution for the dynamic boundaries. (B)

Normalized Hi-C interaction maps are shown for regions with dynamic TAD boundaries in monocytes (top), Day 1 (middle), and Day6 macrophages (bottom). (C) Bar plot of CTCF binding sites distribution relative to TAD boundary. (D) Box plot of CTCF occupancy ratio in Day6 macrophages over monocytes for all TAD boundaries and dynamic TAD boundaries. (E) Distribution of TAD score: the ratio of intra-TAD interactions versus all cis interactions of a TAD, calculated at 40kb resolution. (F) TAD score in Day6 macrophages (y-axis) plotted against TAD score in monocytes (x-axis). All the TADs were sorted by the ratio of their TAD score in monocytes and macrophages and binned into 10 groups. (G) Heatmaps showing the enrichment of several chromatin features (epigenetic modifications, accessibility, CTCF binding, A/B switching) in the TADs in each of the 10 bins.

REFERENCES

1. Italiani P, Italiani P, Boraschi D, Boraschi D. From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. Front Immunol. Frontiers; 2014 Oct 17;5(6):11.
2. Xue J, Schmidt SV, Sander J, Draffehn A, Krebs W, Quester I, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. Immunity. 2014 Feb 20;40(2):274–88.
3. Amit I, Winter DR, Jung S. The role of the local environment and epigenetics in shaping macrophage identity and their effect on tissue homeostasis. Nature Publishing Group. Nature Publishing Group; 2015 Dec 17;17(1):18–25.
4. Quintin J, Saeed S, Martens JHA, Giamarellos-Bourboulis EJ, Ifrim DC, Logie C, et al. Candida albicans infection affords protection against reinfection via functional reprogramming of monocytes. Cell Host Microbe. 2012 Aug 16;12(2):223–32.
5. Saeed S, Quintin J, Kerstens HHD, Rao NA, Aghajanirefah A, Matarese F, et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science. 2014 Sep 26;345(6204):1251086.
6. Pham TH, Benner C, Lichtinger M, Schwarzfischer L, Hu Y, Andreesen R, et al. Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states. Blood. 2012 Jun 14;119(24):e161–71.
7. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular Cell. 2010 May 28;38(4):576–89.
8. Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity. 2010 Mar 26;32(3):317–28.
9. Lavin Y, Winter D, Blecher-Gonen R, David E, Keren-Shaul H, Merad M, et al. Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment. CELL. Elsevier Inc; 2014 Dec 4;159(6):1312–26.
10. Ostuni R, Barozzi I, Termanini A, Bonifacio S, Curina A, Ghisletti S, et al. Latent enhancers activated by stimulation in differentiated cells. CELL. 2013 Jan 17;152(1-2):157–71.
11. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. American Association for the Advancement of Science; 2009 Oct 9;326(5950):289–93.
12. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015 Feb 19;518(7539):331–6.
13. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012 May 17;485(7398):376–80.
14. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012 May 17;485(7398):381–5.

15. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet. Nature Publishing Group; 2015 May 4.
16. Joshi O, Wang S-Y, Kuznetsova T, Atlasi Y, Peng T, Fabre PJ, et al. Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell Stem Cell. 2015 Dec 3;17(6):748–57.
17. Criscione SW, De Cecco M, Siranosian B, Zhang Y, Kreiling JA, Sedivy JM, et al. Reorganization of chromosome architecture in replicative cellular senescence. Sci Adv. 2016 Feb;2(2):e1500882.
18. Chubb JR, Boyle S, Perry P, Bickmore WA. Chromatin motion is constrained by association with nuclear compartments in human cells. Curr Biol. 2002 Mar 19;12(6):439–45.
19. Marshall WF, Straight A, Marko JF, Swedlow J, Dernburg A, Belmont A, et al. Interphase chromosomes undergo constrained diffusional motion in living cells. Curr Biol. 1997 Dec 1;7(12):930–9.
20. Kumaran RI, Spector DL. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J Cell Biol. 2008 Jan 14;180(1):51–65.
21. Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong C-T, et al. Widespread Rearrangement of 3D Chromatin Organization Underlies Polycomb-Mediated Stress-Induced Silencing. Molecular Cell. 2015 Apr 16;58(2):216–31.
22. Taberlay PC, Achinger-Kawecka J, Lun ATL, Buske FA, Sabir K, Gould CM, et al. Three-dimensional disorganization of the cancer genome occurs coincident with long-range genetic and epigenetic alterations. Genome Research. Cold Spring Harbor Lab; 2016 Jun;26(6):719–31.
23. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science. American Association for the Advancement of Science; 2015 Feb 27;347(6225):1017–21.
24. Chandra T, Ewels PA, Schoenfelder S, Furlan-Magaril M, Wingett SW, Kirschner K, et al. Global Reorganization of the Nuclear Landscape in Senescent Cells. Cell Rep. 2015 Jan 28;10(4):471–83.
25. Krijger PHL, Di Stefano B, de Wit E, Limone F, van Oevelen C, de Laat W, et al. Cell-of-Origin-Specific 3D Genome Structure Acquired during Somatic Cell Reprogramming. Cell Stem Cell. 2016 Mar 8.
26. Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ, Bickmore WA. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science. American Association for the Advancement of Science; 2014 Dec 5;346(6214):1238–42.
27. Kuznetsova T, Wang S-Y, Rao NA, Mandoli A, Martens JHA, Rother N, et al. Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization. Genome Biol. BioMed Central Ltd; 2015;16(1):264.
28. Wijchers PJ, Krijger PHL, Geeven G, Zhu Y, Denker A, Verstegen MJAM, et al. Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. Molecular Cell. 2016 Feb 4;61(3):461–73.
29. Schmidt SV, Krebs W, Ulas T, Xue J, Baßler K, Günther P, et al. The transcriptional regulator network of human inflammatory macrophages is defined by open chromatin. Cell Res. Nature Publishing Group; 2016 Jan 5.

30. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature. 2008 Feb 14;451(7180):796–801.
31. Zuin J, Dixon JR, van der Reijden MIJA, Ye Z, Kolovos P, Brouwer RWW, et al. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. Proc Natl Acad Sci USA. National Acad Sciences; 2014 Jan 21;111(3):996–1001.
32. Van Bortle K, Nichols MH, Li L, Ong C-T, Takenaka N, Qin ZS, et al. Insulator function and topological domain border strength scale with architectural protein occupancy. Genome Biol. BioMed Central; 2014;15(6):R82.
33. Denissov S, van Driel M, Voit R, Hekkelman M, Hulsen T, Hernandez N, et al. Identification of novel functional TBP-binding sites and general factor repertoires. The EMBO Journal. 2007 Feb 21;26(4):944–54.
34. Rao NAS, McCalman MT, Moulos P, Francoijs K-J, Chatziioannou A, Kolisis FN, et al. Coactivation of GR and NFKB alters the repertoire of their binding sites and target genes. Genome Research. 2011 Sep;21(9):1404–16.
35. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
36. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Meth. 2012 Oct;9(10):999–1003.

Chapter 6

Discussion

Thanks to the rapid evolution of technologies and the sustained investigations from researchers, a large variety of regulatory elements have been discovered in the genome and their roles in gene regulation have been continuously deciphered. The activities of these elements are found tightly connected with local chromatin configuration. Such local chromatin configuration is often called epigenetic memory or epigenetic dynamics.

In eukaryotic cells, DNA winds around histone proteins forming the basic structural units of chromatin called nucleosomes. Both DNA and histone proteins play a role in epigenetic memory. The histone proteins are especially important, because they can be decorated with various modifications and are deeply associated with the accessibility of the DNA. Such diversity of local chromatin configuration makes epigenetic memory important in regulating gene expression and defining the properties and behavior of the cells. Genome-wide profiling of histone modifications, DNA accessibility and DNA modifications has become a robust and efficient approach to identify functional regulatory elements in distinct cell states or cell types. We studied three different cell systems that mimic the crosstalk in inflammation (**Chapter 2**), the transition between two states of pluripotency (**Chapter 3**), and the differentiation from monocyte to macrophage (**Chapter 4 and 5**), to investigate the impacts of epigenetic memory on gene regulation and to understand the underlying mechanisms that direct the cellular changes.

Impacts of epigenetic memory on gene expression

The epigenetic memory in cells is inheritable, and establishes corresponding stable gene expression patterns. Eukaryotic organisms or at least metazoan utilize epigenetic memory to respond in particular to environmental signals or stimuli over different time scales (D'Urso and Brickner, 2014). Across generations of an organism, such epigenetic memory is also known as genomic imprinting by which the epigenetic marks for the imprinted genes are inherited from the parents whereas the expression patterns of the imprinted genes are maintained in the children (Peters, 2014). Within the same generation of an organism, epigenetic memory defines cell identities and potentials for differentiation during development (Cantone and Fisher, 2013), and establish the abilities of the cells or the next generations of the same cells to respond to the environmental stimuli similar to their previous experiences (Bonasio et al., 2010).

Our studies mainly focused on the relatively short timescale, as it can help us to dissect how cells respond, survive and adapt during environmental changes. All of the three cell systems were manipulated by changing the environment of the cells. In **Chapter 2**, the patterns of epigenetic dynamics together with DNA motif analysis enabled us to further classify

the binding sites of signal-dependent nuclear receptors GR and p65 into constitutive and induced groups that display distinctive regulatory manners. In **Chapter 3**, various epigenetic features at some bivalent promoters were used to determine the differences between the ELRI (Extremely Long-Range Promoter-Promoter Interaction) promoters and non-ELRI promoters. In **Chapter 4**, time-resolved epigenomes and transcriptomes together unveiled the early events following monocyte exposure to LPS or BG and how the resulting epigenetic landscapes determine the function of tolerized and trained macrophages. In **Chapter 5,** we integrated the time-resolved epigenomes and transcriptomes with the chromatin interaction maps from a Capture Hi-C approach to investigate the relationships between epigenetic dynamics and 3D genomic organization during monocyte-to-macrophage differentiation and immunological training. Our studies demonstrated that even stimuli as short as a few hours are able to significantly alter the epigenetic landscape at a collection of genomic regions. Moreover, these epigenetic alterations are strongly associated with gene regulation and 3D genomic organization, and confer to the cells the abilities to respond and survive when they experience similar environmental changes.

Genome-wide study of epigenetic memory extended our understandings of transcriptional regulation beyond protein-coding regions, and allowed the systematic identification of a myriad of potentially functional genomic regions. This approach was initially used to annotate and interpret regulatory DNA elements including enhancers and promoters in a single cell type, or compare them with the elements in other cell types (Chen et al., 2016; Consortium, 2012; Ernst et al., 2011; Roadmap Epigenomics Consortium et al., 2015; Stunnenberg et al., 2016). In our studies, we deployed this approach to track the chromatin changes during cellular responses and differentiation at the epigenetic level. We observed that the epigenetic landscape at promoter regions are largely unchanged, in contrast, epigenetic dynamics occur remarkably more often at enhancer regions especially the active marks such as DHS, H3K27ac, and H3K4me1. In addition, the potential target genes of these dynamic enhancers showed high enrichment in treatment or stimuli related gene ontology terms, and the expression of these gene is highly correlated with the active marks at the enhancers. This implies that enhancers play a vital role in gene regulation over short timescales.

It is well established that an enhancer complies its function via the binding of TFs and other chromatin regulators of chromatin. The TF-binding DNA motifs allow us to predict the candidate TFs that directly bind to the DNA sequence of an enhancer. Moreover, the enrichment of DNA motifs at regions with epigenetic dynamics gives a hint for the master TFs in the processes. Using this approach in **Chapter 4**, we predicted EGR2 and MITF as the key regulators in monocyte-to-macrophage differentiation. These master players can be further examined by ChIP-seq using the antibodies against them or by functional studies such as genomic edit. In conclusion, the hypothesis-free epigenomic and transcriptomic approach

is a powerful and efficient tool in deciphering cell systems and importantly, it generates a number of testable hypotheses.

Linking enhancers to their target genes

Identifying the target genes of enhancers is not a straightforward task, because an enhancer may regulate multiple genes and is frequently (very) distal from its target genes. Three approaches have been used to predict the linkages between enhancers and genes. Firstly, the simplest approach is to assign an enhancer to its nearest transcription start site, but it neglects the fact that enhancers do act over long genomic distances. Secondly, a widely used approach is using the information from epigenomic profiling to correlate the changes in DNA accessibility, histone modifications, or DNA methylation levels with gene expression across different cell types or cellular conditions (Andersson et al., 2014; Ernst et al., 2011; He et al., 2014; Sheffield et al., 2013; Shen et al., 2012; Whalen et al., 2016; Yao et al., 2015). Thirdly, 3C-based techniques are recently being used to predict the linkages by measuring chromatin physical contacts (Jin et al., 2013; Li et al., 2012; Rao et al., 2014; Sanyal et al., 2012; Tang et al., 2015; Zhang et al., 2013). The approach assumes that enhancers make more contacts with functionally relevant regions than with other genomic regions. In **Chapter 2**, **3**, and **4**, we deployed various 3C-based techniques. Our observations are in line with other studies in that either epigenomic-based or 3C-based approaches predict more than one target gene for most enhancers whereas less than half of the enhancers are linked to the nearest TSS (Jin et al., 2013; Li et al., 2012; Sanyal et al., 2012; Schoenfelder et al., 2015a; Zhang et al., 2013). Taken together, epigenomic-based and 3C-based approaches provide more reasonable and comprehensive information about the pairing of regulatory DNA elements and genes within the cell systems.

The epigenomic-based approach to determine the respective targets relies on statistical associations. A few ChIP-seq and RNA-seq experiments are sufficient to provide the essential information to make the prediction for a cell system of interests. However, all of the association methods only provide predictions of putative target genes. Of note, these methods cannot predict the linkages between the regions that are separated by more than several million base-pairs. Because a limit for genomic distance is usually included in the prediction, otherwise, the entire genome is too large and too many regions could be predicted as the targets for individual regions. On the other hand, the 3C-based approach does not have this limitation, however, the technical noise may reduce their power especially when the two genomic regions are (too) far apart.

To answer whether 3C-based techniques are a suitable approach to identify enhancer-promoter linkages, we need to go back to a fundamental question that is how enhancers regulate promoters. To answer it, two non-exclusive mechanistic models have been proposed: the 'tracking' model and the 'looping' model (Blackwood and Kadonaga, 1998). In the 'tracking' model, the transcriptional machinery containing RNA polymerase II (RNAPII) travels over the DNA between an enhancer-promoter pair; whereas in the 'looping' model, an enhancer directly interacts with its target promoters by physical contacts mediated by protein–protein interactions. The 'tracking' model has been shown to exist in sporadic cases in which the enhancers are located only a few kilo base-pairs upstream of the TSSs indicating this model may only apply to the enhancers of close proximity (Hatzis and Talianidis, 2002; Wang et al., 2005). The 'looping' model has been largely supported, because reduction of loop-associated proteins (Ren et al., 2011; Song et al., 2007), interruption of looping by altered insulators (Guo et al., 2015; Lupiáñez et al., 2015; Narendra et al., 2015; Nora et al., 2012; Ong and Corces, 2014) or genomic duplication (Franke et al., 2016) all affect transcriptional outcome. Therefore, using 3C-based techniques seem to be a reasonable approach to identify enhancer-promoter linkages, as these techniques capture the physical contacts between separated genomic regions.

Nevertheless, experimental validation is essential to confirm the predicted enhancer–promoter linkages, and in turn the validation result can also be used to improve the prediction. However, to date, relatively few enhancer–promoter pairs have been experimentally examined (Visel et al., 2007), whereas systematic validation experiments have been performed only in *Drosophila* but not in mammalian cells to support the predicted enhancer-promoter linkages (Kvon et al., 2014).

The frequency of long-range chromatin communication

While the 'looping' model proposes the physical interactions between enhancers and promoters via their associated proteins, it also implicates that the functional chromatin communications such as enhancer-promoter contacts are likely more frequent than with other genomic regions. Since genome-wide 3C-based techniques detect population-averaged crosslinking probabilities of chromatin, they provide a systematic way to estimate and compare the relative frequencies of physical contacts of chromatin between different genomic regions in a global way. It is shown that in pluripotent stem cells enhancers and promoters occupied by the pluripotency factors display higher relative frequencies of intrachromosomal and interchromosomal interactions, suggesting that the cell-type-specific transcription factors may play a key role in shaping genomic 3D architecture (de Wit et al., 2013). However, this observation was based on interaction maps of low resolution (hundreds

of kilo base-pairs) which makes it not easy to determine the specificities of enhancer-promoter contacts. **In Chapter 3 and 5**, our Capture-HiC experiments enabled us to perform a genome-wide detection of chromatin interactions from regions of interest at a resolution of a few kilo base-pairs. We observed that the interacting regions are highly enriched in the distal regulatory DNA elements of the corresponding cell types. This is in line with a deeply sequenced Hi-C study (Rao et al., 2014) and promoter Capture Hi-C studies (Javierre et al., 2016; Mifsud et al., 2015; Schoenfelder et al., 2015b). In conclusion, high relative frequency of physical contacts to promoters is a key attribute of enhancers, and likely plays a role in exerting and modulating enhancer function.

However, the existence of physical contacts between a distal regulatory DNA element and a promoter does not automatically guarantee that the element is actually involved in regulating the expression of that gene. Indeed, in **Chapter 2**, 4C-seq experiments showed that the chromatin interactions between inducible enhancers and target genes displayed significantly higher signals than background even before treatment, while these interactions were almost undetectable in P300 ChIA-PET experiments under the same cell condition. This observation as well as some other studies (Ghavi-Helm et al., 2014; Jin et al., 2013; de Wit et al., 2013) suggest that the interactions between distal regulatory DNA elements and promoters are likely pre-existing. It also implicates that without the required TFs, the enhancer is likely not activated for regulating the interacting gene. Interestingly, upon the treatment with TA and TNFa, the same interactions became detectable in CHIA-PET experiments and showed moderately but significantly higher signals in 4C-seq experiments compared with before treatment. In contrast to 4C-seq, ChIA-PET technique is designed to only capture the chromatin interactions mediated by a protein of interest. Hence, our study suggests that both chromatin physical contacts and TFs contribute to exerting enhancer function and regulating gene expression in which TFs may reinforce the physical contacts between the active enhancers and promoters.

A different way of assessing enhancer-promoter contacts is by DNA fluorescent in situ hybridization (DNA FISH). In contrast to 3C-based techniques, DNA FISH measures physical distances of separated genomic loci in the 3D space of the nucleus. It is also able to depict the distribution of these distances within a chosen cell population. This technique is limited by the number of genomic loci that can be simultaneously monitored. Recently, an improved technique called HIPMap introduced a high-precision, high-throughput, automated pipeline for mapping spatial locations of genomic regions at much larger scale (Shachar et al., 2015). 3D DNA FISH is often used to validate and support the chromatin interactions identified by 3C-based experiments. The majority of the studies suggest that DNA FISH and 3C-based techniques give concordant results over a wide range of genomic and spatial distances (Andersson et al., 2014; Crane et al., 2015; Giorgetti et al., 2014; Kalhor et al., 2012; Li et

al., 2012; Rao et al., 2014), yet discordant result has been reported in a study in which the discrepancy is likely due to the changes in chromatin composition (Williamson et al., 2014). DNA FISH and 3C-based techniques provide intrinsically different but complementary information (Giorgetti and Heard, 2016). In **Chapter 3**, we also deployed DNA FISH in an attempt to validate a selection of identified extremely long-range interactions (ELRIs), however, we were not able to observe significant difference in the distributions of the distances between the interacting loci pairs. This result might be due to the low accuracy of DNA FISH resulting from the technical issues such as the heating step for base-pairing probe and its target DNA (Giorgetti and Heard, 2016; Solovei et al., 2002).

Long-range chromatin interactions are rare events. That is, an interaction between distant loci or loci on different chromosomes occurs infrequently in a cell, or in other words, such interactions occur only in a small number of the cells in a population at any given time. In contrast, loci that are close on the linear genome interact much more frequently, because a chromosome essentially is a large polymer which constrains the spatial distances between the nearby polymeric subunits. With the data from 3C-based techniques, the absolute frequency of a long-range interaction in a cell type can be estimated by calculating the ratio between the normalized read count for the interaction and the total read count for one of the interacting loci. This ratio for an interaction of more than 100 kb genomic distance is typically less than 0.01, which means that in less than 1% of the time the two loci are close enough to make physical contact, or in other words, the physical contacts of the two loci are occurring in less than 1% cells if we take a snapshot of the same cell population. Such low value is in line with the observation from 3D DNA FISH experiments, that the interacting loci can only be observed being located in close proximity in a few cells (Fabre et al., 2015; Schoenfelder et al., 2015b; Vieux-Rochas et al., 2015). Also, the extremely long-range interactions identified in **Chapter 3** displayed low absolute frequencies in both Capture Hi-C and DNA FISH experiments. Hence, it remains unclear what the contribution of such low frequency of chromatin physical contacts is in gene regulation and how cells utilize it to regulate gene expression.

Finally, it is possible that some promoter-interacting regions identified by 3C-based techniques do not show classical enhancer markers or other gene regulatory signatures yet function as enhancer or other types of regulatory DNA elements (Javierre et al., 2016; Rajagopal et al., 2016). It is also possible that distal regulatory DNA elements impact gene expression not through physical interactions with their target genes but through the eRNAs transcribed from these regions, if the eRNAs have an appropriate lifespan. Because non-coding RNAs have been shown to be able to affect TFs, cofactors or signaling molecules thus indirectly influence gene expression (Li et al., 2016).

High-order structure of 3D genomic organization

The next question is how distal regulatory DNA elements are restricted to physically only or mainly interact with their own target genomic loci. The interactions identified by 3C-based techniques in a given cell type include not only the ones involved in gene regulation, but also the ones in maintaining the overall chromosomal structure. The extra information from 3C-based techniques is especially useful for understanding high-level chromatin organization. With the extra information, 3C-based studies in line with DNA FISH studies revealed that the chromosomal structure is compartmentalized into multiple levels including the territories of individual chromosomes, the chromatin compartments of active and inactive regions at the scale of several Mb, and the topologically associating domains (TADs) of large self-interacting regions at the scale of several hundred kb (Bickmore and van Steensel, 2013; Gibcus and Dekker, 2013; Kalhor et al., 2012). Compartmentalization of chromosomes into structural domains is beneficial for chromatin communication. A domain is a relatively isolated structure, largely restricts interactions to occur between the loci within the domain, and prevents the access from other genomic loci. It also implies that the enhancers in a domain are largely constrained to affect the expression of target genes that also reside in the same domain.

Apart from chromosome territories, each chromosome can be classified into A and B compartments which contain relatively active and inactive chromatin, respectively. During stem cell differentiation (Dixon et al., 2015) and monocyte-to-macrophage differentiation (**Chapter 5**), A and B compartments were observed undergoing extensive switching, indicating that high-level chromatin structure is a cell-type specific signature. Indeed, converting somatic cells into induced pluripotent stem cells (iPSCs) effectively erased the somatic-cell-specific A/B compartment patterns, and established an embryonic stem-cell-like pattern (Krijger et al., 2016). However, only a subset of genes displayed differential expression in different cell types, hence it results in a subtle correlation between A/B compartment switching and the changes in overall patterns of gene expression (Dixon et al., 2015; Krijger et al., 2016). Taken together, the A and B compartments provide the first level of cell-type specific environment for individual genes, but do not suffice to determine the cell-type specific patterns of gene expression.

Each chromosome can be further divided into many topologically associating domains (TADs). In contrast to the high cell-type specificity of A/B compartments, TADs are largely invariant in different cell types (Dixon et al., 2015, also in **Chapter 5**). The switching of A/B compartment status during differentiation typically corresponds to a single or series of TADs, suggesting TADs behave as the basic units of the genome. However, the interaction frequency patterns within individual TADs can vary dramatically depending on the cell types. The changes in interaction frequency are positively correlated with the active marks (such

as DHS and H3K27ac) at the interacting regions and negatively correlated with repressive marks (such as H3K27me3) (Dixon et al., 2015). Hence, the internal interaction patterns of TADs are connected with the deposition of epigenetic marks within the TADs. In **Chapter 2,** ChIA-PET experiments uncovered active genomic region centered interaction networks. P300 containing enhancers and the promoters of their target genes predominantly make contacts within segments of the TADs. It indicates the existence of sub-domains with in TADs, which is supported by a 3D map of the human genome at kilobase resolution (Rao et al., 2014). Hence, the contacts of enhancers are restricted by the orchestrated domain-based hierarchical structure of the chromatin.

A chromatin domain is likely formatted by a chromatin loop between two domain boundaries. CTCF, cohesin, condensin, and other architectural proteins are often found at boundaries, in which the orientation of CTCF-binding site relative to each other is a critical factor in setting up the loop (Acemel et al., 2016; de Wit et al., 2015; Dixon et al., 2012; Gómez-Marín et al., 2015; Rao et al., 2014; Tang et al., 2015; Van Bortle et al., 2014). Inversion of CTCF sites disrupts chromatin loops and consequently leads to the establishment of new chromatin loops that are in agreement with the new orientation (Sanborn et al., 2015). In *Caenorhabditis elegans*, it is shown that condensin plays a role in X chromosome dosage compensation by reshaping the X chromosomes into stronger boundaries with more regular spacing compared with autosomes. In **Chapter 5**, we observed a set of boundaries that display changes in boundary strength. However, the exact role of such changes in gene regulation needs further investigation.

Although the domain structure provides an explanation about how distal regulatory DNA elements are restricted to interact with their target genomic loci, the mechanism of TAD formation remains incomplete and need further experimental supports (Dekker and Mirny, 2016). Also, chromatin contacts do occur between domains and chromosomes (Chapter 3 as an example), especially that some of these contacts have been shown functionally important in gene regulation and development (Zhang et al., 2013). However, the principles behind 3D genomic organization and gene regulation are still far from clear.

Enhancer, disease and epigenetic therapy

Given that enhancers play a vital role in orchestrating differentiation, development and environmental response, it is not surprising that many diseases are associated with enhancers. Genetic variants of enhancer-associated TFs (e.g. GATA factors (Lentjes et al., 2016) and Hox factors (Quinonez and Innis, 2014)) and transcriptional coregulators (e.g. CBP (Iyer et al., 2004)) have long been linked to disease. These variants result in dysfunction of

the corresponding proteins thus likely affect many genes that are regulated by these proteins. Genetic variants of individual enhancer regions can also cause disease by altering enhancer activity and target gene expression (Emison et al., 2005; Gröschel et al., 2014; Kulozik et al., 1988). Although genetic variants are an important component in many diseases, the role of epigenetic mechanisms should not be neglected in disease related to early life experience, diet, lifestyle and other environmental effects, as they are responsible for the integration of environmental cues into individual cells at the epigenetic level. Therefore, epigenetic therapy offers a potential way to directly influence related pathways in disease (Arrowsmith et al., 2012; Kelly et al., 2010; Licht, 2015; Portela and Esteller, 2010; Wouters and Delwel, 2016).

In **Chapter 4**, we 'perturbed' normal monocyte-to-macrophage differentiation by exogenous signals to study innate immune memory, with the best-characterized outcomes being endotoxin tolerance and trained innate immunity (Netea et al., 2016). Exposure to high levels of LPS induces a tolerized macrophage phenotype, which is a major cause of sepsis-associated mortality. In contrast, trained immunity has beneficial effects through priming of macrophages for stronger responses to subsequent infection and can be induced by a variety of MAMPs, such as *C. albicans* (Quintin et al., 2012), Bacille Calmette-Gue´rin (BCG) vaccine (Aaby and Benn, 2012), and BG (Saeed et al., 2014). Monocyte exposure to LPS induces and activates a group of enhancers that are associated with an inflammation response, showing discordance in time with accumulation of H3K4me1 while remaining at higher levels in LPS-Mfs compared to naive-Mfs and BG-Mfs. An epigenetic drug, bromodomain and extraterminal domain family inhibitor 151 (IBET151), was shown to be able to disrupt chromatin complexes which are responsible for the expression of key inflammatory genes in activated macrophages, by 'mimicking' acetylated histones (Nicodeme et al., 2010). However, IBET151 was only effective in preventing tolerance when used to block the LPS-induced response via co-treatment with LPS, but was not able to reverse tolerance when administered after LPS. On the other hand, BG treatment after LPS exposure was able to reverse LPS-induced tolerance and to reinstate normal levels of cytokine release in so-called rescue-Mf's. Further investigation of rescue-Mf's at the histone modification level revealed that BG recovers H3K27ac at regions that are silent in LPS-Mf's. This observation further supported the notion that the molecular mechanisms of chromatin remodeling that are induced by BG remain after the initial LPS response, however will be largely blocked by IBET151. Our findings show that the innate immune 'training stimulus' b-glucan outperforms the inflammation blocker IBET151 in reversal of macrophage tolerance. This is an important step towards understanding how the tolerized phenotype can be reversed in sepsis patients and ultimately provides a framework for future therapeutic developments in epigenetic therapy.

Back to genetics

While epigenetic memory shapes cell types and modulate cell functions, individuals do have genetic variations that may cause disease or increase disease susceptibility. Given that 3D genomic organization plays a role in gene regulation, it is interesting to investigate the relationships between genetic variations, epigenetic memory, 3D genomic organization, and gene expression patterns.

Genome wide association studies (GWAS) have identified thousands of single-nucleotide polymorphisms (SNPs) that are associated with particular diseases. Most SNPs are located in non-coding regions (Freedman et al., 2011), in which many of these regions are regulatory DNA elements such as enhancers, suggesting that many SNPs affect gene expression that lead to an increased risk of certain disease (Farh et al., 2015; Gjoneska et al., 2015; Maurano et al., 2012; Roadmap Epigenomics Consortium et al., 2015). This is further supported by the identification of expression quantitative trait loci (eQTL) and histone quantitative trait loci (hQTL) which alter the local chromatin state and enhancer activity (Astle et al., 2016; Chen et al., 2016; Grubert et al., 2015; Kasowski et al., 2013; Kilpinen et al., 2013; McVicker et al., 2013; Waszak et al., 2015). However, QTLs represent a mixture of direct and indirect gene regulation that often cannot be easily distinguished due to the unknown enhancer-promoter linkages. Hence, integration of GWAS SNPs and QTLs with chromatin interaction map provide a good approach to interpret non-coding genetic variation (Javierre et al., 2016).

Large-scale genetic variations are also linked with disease. It has been shown that the patterns of chromosomal rearrangements and translocations are highly correlated with the spatial proximity by the combined analyses of Hi-C and High-throughput genome-wide translocation sequencing (HTGTS) (Zhang et al., 2012). Analyzing patient cells and genetically modified mice with 3C-based methods further showed that deletions, inversions, or duplications can elicit pathogenic effects by changing the higher-order chromatin structure, depending on the size and position of the genetic variations (Franke et al., 2016; Lupiáñez et al., 2015).

3D genomic organization and evolution

Given the significant contribution of 3D genomic organization in gene regulation, it is also important to consider whether and how 3D genomic organization influence genome evolution.

The investigations in a large range of species including bacteria, yeast, plants, and mammals suggest that the form of TAD-like structures seems to be a common mechanism that cells

organize their genomes in 3D space (Hsieh et al., 2015; Le et al., 2013; Mizuguchi et al., 2014; Wang et al., 2015). Interestingly, apart from the high similarity between the linear genome sequences of mammals, syntenic genomic regions also display remarkable conservation in TAD organization (Dixon et al., 2012; Vietri Rudan et al., 2015). This conservation is likely dependent on the conserved CTCF binding sites, which together with cohesin determine the TAD boundaries. The conservation of 3D genomic organization extends beyond the TADs. Evolutionary related genomic regions that are distant in the human genome but are adjacent in the mouse genome were shown to have significantly more chromatin contacts than expected (Véron et al., 2011). In our study in **Chapter 3,** the extremely long-range promoter-promoter Interaction (ELRI) loci identified in primed mouse pluripotent stem cells largely overlap with the promoters of many homeobox genes. Interestingly, similar long-range chromatin interactions have also been observed in *Drosophila* (Bantignies et al., 2011), whereas these homeobox genes are believed to be generated by genome duplication during metazoan evolution (Garcia-Fernàndez, 2005). Therefore, the 3D genomic organization is conserved at various levels.

3D genomic organization is likely also linked with the morphological alterations during evolution. In the mouse genome, the *HoxD* gene cluster is flanked by two TADs. With this configuration, the *HoxD* genes can be regulated by the regulatory DNA elements in the anterior and posterior TADs on demand to form the appendages of the body such as limbs (Andrey et al., 2013). This seems to be conserved in vertebrates, however, the *Hox* cluster in the embryos of an invertebrate chordate called amphioxus is organized into a single TAD that includes long-range contacts mostly from the anterior side (Acemel et al., 2016). Interestingly, most of the regulatory elements within the vertebrate-specific posterior TAD are required for the formation of vertebrate-specific appendages (Lonfat et al., 2014; Montavon et al., 2011). This example indicates that 3D genomic organization may facilitate the incorporation of newly emerged regulatory input to the existing regulatory information and contribute to the morphological alterations during evolution.

In summary, 3D genomic organization is associated with evolution and may also contribute to the morphological alterations during evolution. However, further investigations in a larger number of species are required to infer the general roles of 3D genomic organization in evolution.

Perspective

Our view of the roles of distal regulatory DNA elements and the principles guiding chromatin organization are becoming greatly expanded, and have led us to a more comprehensive

understanding of the genome and the cell systems. However, further studies are needed in many aspects. Firstly, what are the common features of active enhancers? Current studies identified a large number of enhancers based on epigenetic marks. However, it has been shown that only a subset of the H3K27ac marked enhancers in a cell type are validated to be functionally active using transgenic mouse assays (Nord et al., 2013). Thus, additional experiments are needed to gather the actual activities of enhancers in individual cellular contexts. Secondly, what is the nature of enhancer-promoter communication? This includes broad questions such as how enhancers regulate genes, how enhancers find their target genes, how frequently enhancers make contact with their targets, how the contact frequencies are translated into transcription, and what are the contributions of individual enhancers in a complex enhancer-promoter interaction network. Thirdly, many underlying mechanisms remain unclear, including the formation process of chromatin loops, the dynamics of chromatin physical contacts, and the cell-to-cell variability in chromosome folding and its role in gene regulation (Dekker and Mirny, 2016).

To address these questions, further supports from experimental evidence are needed. However, our current understanding of the activities, functions, and corresponding phenotypes of regulatory DNA elements largely relies on time consuming onebyone experimental methods. The promotion of high throughput functional assays such as massively parallel reporter assays (Kvon et al., 2014) and genomeediting screenings (Shalem et al., 2015) would lead us to a more comprehension view of regulatory DNA elements. Our understanding of chromatin organization and its impacts on gene regulation is also largely limited, by current population-averaged measurement. A deeper understanding into the dynamics of chromatin organization and underlying mechanisms of chromatin physical contact will provide further mechanistic insights into gene regulation and its variability in individual cells. This may require improvement in experimental methods at single-cell level, and, which is even better, real-time in live cells. Using live-imaging methods and quantitative analysis, a recent study provides a nice start in understanding the associations between enhancer and transcriptional bursting (Fukaya et al., 2016). Ultimately, the up-coming fundamental understanding of the regulatory code and grammar will be beneficial for guiding the therapy of disease.

REFERENCES

1. Aaby, P., and Benn, C.S. (2012). Saving lives by training innate immunity with bacille Calmette-Guérin vaccine. Proc. Natl. Acad. Sci. *109*, 17317–17318.
2. Acemel, R.D., Tena, J.J., Irastorza-Azcarate, I., Marlétaz, F., Gómez-Marín, C., de la Calle-Mustienes, E., Bertrand, S., Diaz, S.G., Aldea, D., Aury, J.-M., et al. (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. Nat. Genet. *48*, 336–341.
3. Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.
4. Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., Spitz, F., and Duboule, D. (2013). A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. Science *340*, 1234167.
5. Arrowsmith, C.H., Bountra, C., Fish, P.V., Lee, K., and Schapira, M. (2012). Epigenetic protein families: a new frontier for drug discovery. Nat. Rev. Drug Discov. *11*, 384–400.
6. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell *167*, 1415–1429.e19.
7. Bantignies, F., Roure, V., Comet, I., Leblanc, B., Schuettengruber, B., Bonnet, J., Tixier, V., Mas, A., and Cavalli, G. (2011). Polycomb-Dependent Regulatory Contacts between Distant Hox Loci in Drosophila. Cell *144*, 214–226.
8. Bickmore, W.A., and van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. Cell *152*, 1270–1284.
9. Blackwood, E.M., and Kadonaga, J.T. (1998). Going the Distance: A Current View of Enhancer Action. Science *281*, 60–63.
10. Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular Signals of Epigenetic States. Science *330*, 612–616.
11. Cantone, I., and Fisher, A.G. (2013). Epigenetic programming and reprogramming during development. Nat. Struct. Mol. Biol. *20*, 282–289.
12. Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yan, Y., Kundu, K., Ecker, S., et al. (2016). Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. Cell *167*, 1398–1414.e24.
13. Consortium, T.E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.
14. Crane, E., Bian, Q., McCord, R.P., Lajoie, B.R., Wheeler, B.S., Ralston, E.J., Uzawa, S., Dekker, J., and Meyer, B.J. (2015). Condensin-driven remodelling of X chromosome topology during dosage compensation. Nature *523*, 240–244.

15. Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110–1121.
16. de Wit, E., Vos, E.S.M., Holwerda, S.J.B., Valdes-Quezada, C., Verstegen, M.J.A.M., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H.L., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Mol. Cell *60*, 676–684.
17. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.
18. Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.
19. D'Urso, A., and Brickner, J.H. (2014). Mechanisms of epigenetic memory. Trends Genet. *30*, 230–236.
20. Emison, E.S., McCallion, A.S., Kashuk, C.S., Bush, R.T., Grice, E., Lin, S., Portnoy, M.E., Cutler, D.J., Green, E.D., and Chakravarti, A. (2005). A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. Nature *434*, 857–863.
21. Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shoresh, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. Nature *473*, 43–49.
22. Fabre, P.J., Benke, A., Joye, E., Huynh, T.H.N., Manley, S., and Duboule, D. (2015). Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. Proc. Natl. Acad. Sci. *112*, 13964–13969.
23. Farh, K.K.-H., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W.J., Beik, S., Shoresh, N., Whitton, H., Ryan, R.J.H., Shishkin, A.A., et al. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature *518*, 337–343.
24. Franke, M., Ibrahim, D.M., Andrey, G., Schwarzer, W., Heinrich, V., Schöpflin, R., Kraft, K., Kempfer, R., Jerković, I., Chan, W.-L., et al. (2016). Formation of new chromatin domains determines pathogenicity of genomic duplications. Nature *538*, 265–269.
25. Freedman, M.L., Monteiro, A.N.A., Gayther, S.A., Coetzee, G.A., Risch, A., Plass, C., Casey, G., De Biasi, M., Carlson, C., Duggan, D., et al. (2011). Principles for the post-GWAS functional characterization of cancer risk loci. Nat. Genet. *43*, 513–518.
26. Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. Cell *166*, 358–368.
27. Garcia-Fernàndez, J. (2005). The genesis and evolution of homeobox gene clusters. Nat. Rev. Genet. *6*, 881–892.
28. Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W., and Furlong, E.E.M. (2014). Enhancer loops appear stable during development and are associated with paused polymerase. Nature *512*, 96–100.
29. Gibcus, J.H., and Dekker, J. (2013). The Hierarchy of the 3D Genome. Mol. Cell *49*, 773–782.
30. Giorgetti, L., and Heard, E. (2016). Closing the loop: 3C versus DNA FISH. Genome Biol. *17*, 215.

31. Giorgetti, L., Galupa, R., Nora, E.P., Piolot, T., Lam, F., Dekker, J., Tiana, G., and Heard, E. (2014). Predictive Polymer Modeling Reveals Coupled Fluctuations in Chromosome Conformation and Transcription. Cell *157*, 950–963.
32. Gjoneska, E., Pfenning, A.R., Mathys, H., Quon, G., Kundaje, A., Tsai, L.-H., and Kellis, M. (2015). Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer/'s disease. Nature *518*, 365–369.
33. Gómez-Marín, C., Tena, J.J., Acemel, R.D., López-Mayorga, M., Naranjo, S., Calle-Mustienes, E. de la, Maeso, I., Beccari, L., Aneas, I., Vielmas, E., et al. (2015). Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. Proc. Natl. Acad. Sci. *112*, 7542–7547.
34. Gröschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A.M., Erpelinck, C., van der Velden, V.H.J., Havermans, M., Avellino, R., van Lom, K., et al. (2014). A Single Oncogenic Enhancer Rearrangement Causes Concomitant EVI1 and GATA2 Deregulation in Leukemia. Cell *157*, 369–381.
35. Grubert, F., Zaugg, J.B., Kasowski, M., Ursu, O., Spacek, D.V., Martin, A.R., Greenside, P., Srivas, R., Phanstiel, D.H., Pekowska, A., et al. (2015). Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions. Cell *162*, 1051–1065.
36. Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell *162*, 900–910.
37. Hatzis, P., and Talianidis, I. (2002). Dynamics of Enhancer-Promoter Communication during Differentiation-Induced Gene Activation. Mol. Cell *10*, 1467–1477.
38. He, B., Chen, C., Teng, L., and Tan, K. (2014). Global view of enhancer–promoter interactome in human cells. Proc. Natl. Acad. Sci. *111*, E2191–E2199.
39. Hsieh, T.-H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108–119.
40. Iyer, N.G., Özdag, H., and Caldas, C. (2004). p300/CBP and cancer. Oncogene *23*, 4225–4231.
41. Javierre, B.M., Burren, O.S., Wilder, S.P., Kreuzhuber, R., Hill, S.M., Sewitz, S., Cairns, J., Wingett, S.W., Várnai, C., Thiecke, M.J., et al. (2016). Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell *167*, 1369–1384.e19.
42. Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.-A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.
43. Kalhor, R., Tjong, H., Jayathilaka, N., Alber, F., and Chen, L. (2012). Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat. Biotechnol. *30*, 90–98.
44. Kasowski, M., Kyriazopoulou-Panagiotopoulou, S., Grubert, F., Zaugg, J.B., Kundaje, A., Liu, Y., Boyle, A.P., Zhang, Q.C., Zakharia, F., Spacek, D.V., et al. (2013). Extensive Variation in Chromatin States Across Humans. Science *342*, 750–752.

45. Kelly, T.K., De Carvalho, D.D., and Jones, P.A. (2010). Epigenetic modifications as therapeutic targets. Nat. Biotechnol. *28*, 1069–1078.
46. Kilpinen, H., Waszak, S.M., Gschwind, A.R., Raghav, S.K., Witwicki, R.M., Orioli, A., Migliavacca, E., Wiederkehr, M., Gutierrez-Arcelus, M., Panousis, N.I., et al. (2013). Coordinated Effects of Sequence Variation on DNA Binding, Chromatin Structure, and Transcription. Science *342*, 744–747.
47. Krijger, P.H.L., Di Stefano, B., de Wit, E., Limone, F., van Oevelen, C., de Laat, W., and Graf, T. (2016). Cell-of-Origin-Specific 3D Genome Structure Acquired during Somatic Cell Reprogramming. Cell Stem Cell *18*, 597–610.
48. Kulozik, A.E., Kar, B.C., Serjeant, G.R., Serjeant, B.E., and Weatherall, D.J. (1988). The molecular basis of alpha thalassemia in India. Its interaction with the sickle cell gene. Blood *71*, 467–472.
49. Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature *512*, 91–95.
50. Le, T.B.K., Imakaev, M.V., Mirny, L.A., and Laub, M.T. (2013). High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. Science *342*, 731–734.
51. Lentjes, M.H., Niessen, H.E., Akiyama, Y., Bruïne, A.P. de, Melotte, V., and Engeland, M. van (2016). The emerging role of GATA transcription factors in development and disease. Expert Rev. Mol. Med. *18*.
52. Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. Cell *148*, 84–98.
53. Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat. Rev. Genet. *17*, 207–223.
54. Licht, J.D. (2015). DNA Methylation Inhibitors in Cancer Therapy: The Immunity Dimension. Cell *162*, 938–939.
55. Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science *346*, 1004–1006.
56. Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell *161*, 1012–1025.
57. Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science *337*, 1190–1195.
58. McVicker, G., van de Geijn, B., Degner, J.F., Cain, C.E., Banovich, N.E., Raj, A., Lewellen, N., Myrthil, M., Gilad, Y., and Pritchard, J.K. (2013). Identification of genetic variants that affect histone modifications in human cells. Science *342*, 747–749.
59. Mifsud, B., Tavares-Cadete, F., Young, A.N., Sugar, R., Schoenfelder, S., Ferreira, L., Wingett, S.W., Andrews, S., Grey, W., Ewels, P.A., et al. (2015). Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat. Genet. *47*, 598–606.

60. Mizuguchi, T., Fudenberg, G., Mehta, S., Belton, J.-M., Taneja, N., Folco, H.D., FitzGerald, P., Dekker, J., Mirny, L., Barrowman, J., et al. (2014). Cohesin-dependent globules and heterochromatin shape 3D genome architecture in S. pombe. Nature *516*, 432–435.
61. Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A Regulatory Archipelago Controls Hox Genes Transcription in Digits. Cell *147*, 1132–1145.
62. Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science *347*, 1017–1021.
63. Netea, M.G., Joosten, L.A.B., Latz, E., Mills, K.H.G., Natoli, G., Stunnenberg, H.G., O'Neill, L.A.J., and Xavier, R.J. (2016). Trained immunity: A program of innate immune memory in health and disease. Science *352*, aaf1098.
64. Nicodeme, E., Jeffrey, K.L., Schaefer, U., Beinke, S., Dewell, S., Chung, C., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., et al. (2010). Suppression of inflammation by a synthetic histone mimic. Nature *468*, 1119–1123.
65. Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381–385.
66. Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and Pervasive Changes in Genome-wide Enhancer Usage during Mammalian Development. Cell *155*, 1521–1531.
67. Ong, C.-T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat. Rev. Genet. *15*, 234–246.
68. Peters, J. (2014). The role of genomic imprinting in biology and disease: an expanding view. Nat. Rev. Genet. *15*, 517–530.
69. Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nat. Biotechnol. *28*, 1057–1068.
70. Quinonez, S.C., and Innis, J.W. (2014). Human HOX gene disorders. Mol. Genet. Metab. *111*, 4–15.
71. Quintin, J., Saeed, S., Martens, J.H.A., Giamarellos-Bourboulis, E.J., Ifrim, D.C., Logie, C., Jacobs, L., Jansen, T., Kullberg, B.-J., Wijmenga, C., et al. (2012). Candida albicans Infection Affords Protection against Reinfection via Functional Reprogramming of Monocytes. Cell Host Microbe *12*, 223–232.
72. Rajagopal, N., Srinivasan, S., Kooshesh, K., Guo, Y., Edwards, M.D., Banerjee, B., Syed, T., Emons, B.J.M., Gifford, D.K., and Sherwood, R.I. (2016). High-throughput mapping of regulatory DNA. Nat. Biotechnol. *34*, 167–174.
73. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

74. Ren, X., Siegel, R., Kim, U., and Roeder, R.G. (2011). Direct Interactions of OCA-B and TFII-I Regulate Immunoglobulin Heavy-Chain Gene Transcription by Facilitating Enhancer-Promoter Communication. Mol. Cell *42*, 342–355.
75. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.
76. Saeed, S., Quintin, J., Kerstens, H.H.D., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.-C., Ratter, J., Berentsen, K., Ent, M.A. van der, et al. (2014). Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science *345*, 1251086.
77. Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. *112*, E6456–E6465.
78. Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109–113.
79. Schoenfelder, S., Furlan-Magaril, M., Mifsud, B., Tavares-Cadete, F., Sugar, R., Javierre, B.-M., Nagano, T., Katsman, Y., Sakthidevi, M., Wingett, S.W., et al. (2015a). The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. Genome Res. *25*, 582–597.
80. Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.-M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., et al. (2015b). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat. Genet. *47*, 1179–1186.
81. Shachar, S., Voss, T.C., Pegoraro, G., Sciascia, N., and Misteli, T. (2015). Identification of Gene Positioning Factors Using High-Throughput Imaging Mapping. Cell *162*, 911–923.
82. Shalem, O., Sanjana, N.E., and Zhang, F. (2015). High-throughput functional genomics using CRISPR-Cas9. Nat. Rev. Genet. *16*, 299–311.
83. Sheffield, N.C., Thurman, R.E., Song, L., Safi, A., Stamatoyannopoulos, J.A., Lenhard, B., Crawford, G.E., and Furey, T.S. (2013). Patterns of regulatory activity across diverse human cell types predict tissue identity, transcription factor binding, and long-range interactions. Genome Res. *23*, 777–788.
84. Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116–120.
85. Solovei, I., Cavallo, A., Schermelleh, L., Jaunin, F., Scasselati, C., Cmarko, D., Cremer, C., Fakan, S., and Cremer, T. (2002). Spatial Preservation of Nuclear Chromatin Architecture during Three-Dimensional Fluorescence in Situ Hybridization (3D-FISH). Exp. Cell Res. *276*, 10–23.
86. Song, S.-H., Hou, C., and Dean, A. (2007). A Positive Role for NLI/Ldb1 in Long-Range β-Globin Locus Control Region Function. Mol. Cell *28*, 810–822.
87. Stunnenberg, H.G., Abrignani, S., Adams, D., de Almeida, M., Altucci, L., Amin, V., Amit, I., Antonarakis, S.E., Aparicio, S., Arima, T., et al. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell *167*, 1145–1149.

88. Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell *163*, 1611–1627.
89. Van Bortle, K., Nichols, M.H., Li, L., Ong, C.-T., Takenaka, N., Qin, Z.S., and Corces, V.G. (2014). Insulator function and topological domain border strength scale with architectural protein occupancy. Genome Biol. *15*, R82.
90. Véron, A.S., Lemaitre, C., Gautier, C., Lacroix, V., and Sagot, M.-F. (2011). Close 3D proximity of evolutionary breakpoints argues for the notion of spatial synteny. BMC Genomics *12*, 303.
91. Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. Cell Rep. *10*, 1297–1309.
92. Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015). Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. Proc. Natl. Acad. Sci. *112*, 4672–4677.
93. Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. Nucleic Acids Res. *35*, D88–D92.
94. Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., Lanz, C., and Weigel, D. (2015). Genome-wide analysis of local chromatin packing in Arabidopsis thaliana. Genome Res. *25*, 246–256.
95. Wang, Q., Carroll, J.S., and Brown, M. (2005). Spatial and Temporal Recruitment of Androgen Receptor and Its Coactivators Involves Chromosomal Looping and Polymerase Tracking. Mol. Cell *19*, 631–642.
96. Waszak, S.M., Delaneau, O., Gschwind, A.R., Kilpinen, H., Raghav, S.K., Witwicki, R.M., Orioli, A., Wiederkehr, M., Panousis, N.I., Yurovsky, A., et al. (2015). Population Variation and Genetic Control of Modular Chromatin Architecture in Humans. Cell *162*, 1039–1050.
97. Whalen, S., Truty, R.M., and Pollard, K.S. (2016). Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat. Genet. *48*, 488–496.
98. Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R.S., Paquette, D., Dostie, J., and Bickmore, W.A. (2014). Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. Genes Dev. *28*, 2778–2791.
99. de Wit, E., Bouwman, B.A.M., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J.A.M., Krijger, P.H.L., Festuccia, N., Nora, E.P., Welling, M., et al. (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227–231.
100. Wouters, B.J., and Delwel, R. (2016). Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. Blood *127*, 42–52.
101. Yao, L., Shen, H., Laird, P.W., Farnham, P.J., and Berman, B.P. (2015). Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. Genome Biol. *16*, 105.

102. Zhang, Y., McCord, R.P., Ho, Y.-J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial Organization of the Mouse Genome and Its Role in Recurrent Chromosomal Translocations. Cell *148*, 908–921.

103. Zhang, Y., Wong, C.-H., Birnbaum, R.Y., Li, G., Favaro, R., Ngan, C.Y., Lim, J., Tai, E., Poh, H.M., Wong, E., et al. (2013). Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. Nature *504*, 306–310.

Summary

SUMMARY

How a genome governs phenotypic plasticity, has long been the subject of research. Gene regulation has been associated with the activities of regulatory DNA elements, epigenetic memory and 3D genomic organization. Here, we investigated the effects of chemical signals on chromatin re-organization, the differences in the chromatin organization of naïve and primed embryonic stem cells, and epigenetic memories, chromatin reorganization and reversal of tolerance in innate immune cells.

In **Chapter 2**, using circular chromosome conformation capture coupled with next generation sequencing (4C-seq) and high-resolution chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) of P300, we investigated the impact of signal-dependent transcription factors glucocorticoid receptor (GR) and nuclear factor kappa-b (NF-κB) on gene regulation and three-dimensional (3D) chromatin organization. We observed agonist-induced changes in long-range chromatin interactions, and uncovered interconnected enhancer–enhancer hubs spanning up to one megabase. The vast majority of activated glucocorticoid receptor and nuclear factor kappa-b appeared to join pre-existing P300 enhancer hubs without affecting the chromatin conformation. In contrast, binding of the activated transcription factors to loci with their consensus response elements led to the increased formation of an active epigenetic state of enhancers and a significant increase in long-range interactions within pre-existing enhancer networks. De novo enhancers or ligand-responsive enhancer hubs preferentially interacted with ligand-induced genes. We demonstrate that, at a subset of genomic loci, ligand-mediated induction leads to active enhancer formation and an increase in long-range interactions, facilitating efficient regulation of target genes. Therefore, our data suggest an active role of signal-dependent transcription factors in chromatin and long-range interaction remodeling.

In **Chapter 3**, we established Capture Hi-C with target-sequence enrichment of DNase I hypersensitive sites to assess the dynamics of 3D genomic organization during the transition between two states of pluripotency serum-to-2i interconversion of mouse embryonic stem cells. We detected extremely long-range intra- and inter- chromosomal interactions between a small subset of H3K27me3 marked bivalent promoters involving the Hox clusters in serum-grown cells. Notably, these promoter-mediated interactions are not present in 2i ground-state pluripotent mESCs but appear upon their further development into primed-like serum mESCs. Reverting serum mESCs to ground-state 2i mESCs removes these promoter-promoter interactions in a spatiotemporal manner. H3K27me3, which is largely absent at bivalent promoters in ground-state 2i mESCs, is necessary, but not sufficient, to establish these interactions, as confirmed by Capture Hi-C on $Eed^{-/-}$ serum mESCs. Our results implicate

H3K27me3 and PRC2 as critical players in chromatin alteration during priming of ESCs for differentiation.

In **Chapter 4**, we integrated time-resolved epigenomes and transcriptomes to study innate immune memory. We characterized the molecular events involved in LPS stimulated, beta-glucan stimulated and normal monocyte-to-macrophage differentiation. We observed that LPS-treated monocytes fail to accumulate active histone marks at promoter and enhancers of genes in the lipid metabolism and phagocytic pathways. Transcriptional inactivity in response to a second LPS exposure in tolerized macrophages is accompanied by failure to deposit active histone marks at promoters of tolerized genes. In contrast, b-glucan partially reverses the LPS-induced tolerance in vitro. Importantly, ex vivo b-glucan treatment of monocytes from volunteers with experimental endotoxemia reinstates their capacity for cytokine production. Tolerance is reversed at the level of distal element histone modification and transcriptional reactivation of otherwise unresponsive genes. Our findings show that the innate immune ''training stimulus'' beta-glucan can reverse macrophage tolerance ex vivo and move forward in understanding how the tolerized phenotype can be reversed in sepsis patients.

Lastly, in **Chapter 5** we performed Capture Hi-C to investigate the dynamics of 3D genomic organization during monocyte-to-macrophage differentiation and immunological training. The monocyte-to-macrophage differentiation is associated with changes at epigenetic, transcriptional and 3D genomic organization levels. We demonstrate that despite the absence of cell division, this differentiation is associated with higher-order chromatin reorganization, including changes in A/B compartments, strengthening of TADs and their boundaries and a global gain of short-range interactions. While differentiation is associated with profound changes, immunological training does not cause spatial chromatin dynamics, suggesting that differentiation is the main driver of chromatin remodeling.

In summary, our studies demonstrated that even stimuli as short as several hours are able to significantly alter the epigenetic landscape at a collection of genomic regions. Moreover, these epigenetic alterations are strongly associated with gene regulation and 3D genomic organization. Our findings provide important insights into gene regulation at epigenetic and 3D genomic organization level.

Samenvatting

SAMENVATTING

Hoe het genoom met phenotypisch plasticiteit omgaat is al geruime tijd onderwerp van studie. Genregulatie is geassocieerd met de activiteit van regulatoire DNA elementen, het epigenetische geheugen en 3D genoom organisatie. In deze studie hebben we het effect van chemische signalen op de reorganisatie van het chromatine onderzocht, de verschillen in chromatine organisatie in naive stamcellen en op differentiatie-voorbereidde stamcellen, en het epigenetisch geheugen, chromatine reorganisatie en omkering van tolerantie in aangeboren immuuncellen.

In **hoofdstuk 2**, hebben we de impact van de signaal-afhankelijke transcriptiefactoren glucocorticoid receptor (GR) en nuclear factor kappa-b (NF-kB) op genregulatie en de drie-dimensionale (3D) chromatine organisatie onderzocht. We observeerden agonist-geïnduceerde veranderingen in lange-afstands chromatine interacties, en ontdekten enhancer-enhancer interacties die reikten tot over een megabase. Het overgrote deel van de geactiveerde glucocorticoide receptoren en NFkB bleek zich te voegen bij reeds aanwezig P300 enhancer regionen zonder de conformatie van het chromatine te beïnvloeden. Binding van deze geactiveerde transcriptiefactoren op hun reeds bekende specifieke DNA bindingselementen leidde tot een verhoogde epigenetische activatie van enhancers en een significante toename in lange-afstands interacties binnen de reeds aanwezige enhancer netwerken. De novo enhancers of ligand-gevoelige enhancer gebieden interacteerden bij voorkeur met ligand-geïnduceerde genen. Wij demonstreerden dat, in bepaalde delen van het genoom, ligand-gemedieerde inductie leidt tot actieve enhancer formatie en een toename in lange-afstands interacties, dat efficiënte regulatie van target genen faciliteert. Onze data suggereert daarom een actieve rol van signaal-afhankelijke transcriptie factoren in de herstructurering van chromatine en lange-afstands interacties.

In **hoofdstuk 3** hebben Capture Hi-C met verrijking van DNAse I gevoelige gebieden opgezet om de dynamiek van de 3D genoom organisatie tijdens de transitie tussen twee verschillende staten van pluripotentie in kaart te brengen. We detecteerden extreem lange-afstands interacties, inter- en intra-chromosomaal, in een klein gedeelte van de H3K27me3 gemarkeerde bivalente promoters, welke betrokken zijn bij de regulatie van de Hox clusters in serum-gegroeide cellen. In het bijzonder, deze promoter-gemedieerde interacties zijn afwezig in 2i mESCs maar verschijnen tijdens ontwikkeling naar op differentiatie-voorbereidde mESCs. Bij de omkering van 2i naar serum condities verdwijnen deze promoter-promoter interacties in een tijdruimtelijke manier. H3K27me3, dat grotendeels afwezig is op bivalente promoters in 2i mESCs, is vereist, maar niet genoeg, om deze interacties tot stand te brengen, dat bevestigd werd door Capture Hi-C in Eed$^{-/-}$ serum ESCs. Onze resultaten impliceren dat H3K27me3 en PRC2 een kritische spelen bij chromatine veranderingen tijdens het voorbereiden van mESCs op differentiatie.

In **hoofdstuk 4** hebben we het epigenoom en transcriptoom door de tijd gebruikt om het aangeboren immuunsysteem te bestuderen. We hebben de moleculaire gebeurtenissen tijdens LPS-stimulatie, beta-glycaan-stimulatie en normale monocyt-naar-macrofaag differentiatie gekarakteriseerd. We observeerden dat LPS-gestimuleerde monocyten niet in staat waren om actieve histon markeringen op de promoters en enhancers van genen betrokken bij vet metabolisme en phagocytose te accumuleren. Transcriptionele inactiviteit in reactie op een tweede blootstelling aan LPS in tolerante macrofagen gaat gepaard met het onvermogen om actieve histon markeringen af te zetten op de promoters van tolerante genen. Beta-glycaan hief de LPS-gestimuleerde tolerantie daarentegen deels op. Belangrijk, een ex vivo beta-glycaan behandeling van monocyten van vrijwilligers met experimentele endotoxemia herintroduceert hun vermogen om cytokines te produceren. Tolerantie is omgekeerd op het niveau van histon modificaties op distale elementen en transcriptionele reactivatie van anders niet-reactieve genen. Onze bevindingen tonen dat de aangeboren immuunsysteem “training stimulus” beta-glycaan tolerante macrofagen ex vivo kan reactiveren en helpt ons te begrijpen hoe het tolerante fenotype in sepsis patiënten omgekeerd kan worden.

Tenslotte, in **hoofdstuk 5** hebben we Capture Hi-C gebruikt om de dynamiek van 3D genoom herstructurering tijdens monocyt-naar-macrofaag differentiatie en immunologische training te onderzoeken. De differentiatie van monocyt naar macrofaag gaat gepaard met veranderingen op epigenetisch, transcriptioneel en 3D genoom organisatie niveau. We demonstreren dat, ondanks de afwezigheid van celdeling, differentiatie is geassocieerd met chromatine reorganisatie, inclusief A/B compartimenten, het versterken van TADs en hun grenzen en een globale toename van het aantal korte-afstands interacties. Hoewel differentiatie is geassocieerd met grondige veranderingen veroorzaakt immunologische training geen structurele veranderingen in het chromatine, suggererend dat differentiatie de belangrijkste oorzaak van chromatine herstructurering is.

Samenvattend, onze studies tonen aan dat stimuli, reeds met de duur van slechts enkele uren, in staat zijn significante veranderingen in het epigenetische landschap teweeg te brengen. Bovendien zijn deze veranderingen sterk geassocieerd met genregulatie en de 3D organisatie van het genoom. Onze bevindingen verschaffen belangrijke inzichten in genregulatie op het epigenetisch niveau en op het niveau van de 3D structuur van het genoom.

Acknowledgement

ACKNOWLEDGEMENT

When I was in my teenager, I always had a dream that I could become a scientist travelling and studying in Europe in which the modern science had been born. This dream came true after a skype interview with an amiable and energetic old man. That man was my supervisor Henk Stunnenberg.

Thank you, Henk! No words would be good enough to express my gratitude towards you. During the past four years, you have been a great supervisor, kind, supportive, and strict. I will forever be thankful for everything you have done for me. You have given me so much freedom to pursue various projects. Discussing with you on science has always been an enjoyable time, as you have always provided insightful ideas and critical inputs. You also have been a nice friend and like a father to me. I still remember your smiles when the first time I saw you via skype, the first day I come to the lab when you handled a map of Nijmegen to me, and many other times. I felt so sad when you told us that you got colorectal cancer. Right after your surgery, it was a relief that knowing you were doing well while you were eager to discuss about the final piece of our manuscript for revision. It is my great honor to be your student. There are countless memories with you that I will never forget in the rest of my life. I wish you all the best!

I must thank Guoliang. Without you, I would not have had the great opportunity to be in Henk's wonderful lab. I also thank my master's supervisor Hong-Yu. You woke up the little boy inside me, a boy with curiosity about every amazing thing in nature. Thank both of you for your helpful career advice and academic suggestions.

I have had a great time in Henk's lab. People here are genuinely nice, and have help me a lot to survive and stay sane during my PhD journey. I would like to take this opportunity to thank all of you.

I would like to thank my daily supervisors Nagesha and Lusy, and my first-year mentor Kees. I learned a lot of essential knowledge from you. Nagesha, thank you for taking care of me when I literally knew nothing. Lusy, thank you for introducing me to your friends. Both of you made the beginning of my life in an unfamiliar country much easier. Kees, thank you for taking out so much time for me. It was so lucky to have you as my mentor during my first year in the lab. I enjoyed every discussion with you on my project, as you provided me so many objective opinions from a statistician's view.

I also would like to thank all of my great and intelligent colleges and collaborators. As a pure bioinformatician, nothing else but bench biologists are more important to me. Tatyana,

Onkar, Boris, and Ehsan, thank you for all your hard working to generate solid and elegant data! Tatyana and Onkar, thank you for your perseverance when we had quite difficult time that the experiments kept failing. Boris, thank you for letting me be involved in your projects and teaching me so much knowledge of hematology. Thank you, Ehsan, you are such a nice and funny person to work with. It has been a pleasure working with all of you and thank you for your support and care.

I was so lucky to be a member of the bioinformatics group. I have learnt a wide range of knowledge and tons of useful tools from each of you. I know that I could always ask you for advice and opinions when I have any bioinformatics related issue. Simon, thank you for making time to explain every piece of detail you know about motif analysis. Hinri, thank you for sharing your experiences and opinions in various computational pipelines. Kees-Jan, talking with you always relaxed us during coffee breaks, and your tips on shell language and server management have help me a lot. Arjen, thank you for sharing me your knowledge about stem cells and DNA methylation. Martin, you always explain things as clear as you can when I come to you with any bioinformatics related issue. Also, it was really nice that sometimes have a talk with you before going home from work. Thank you, Philip, I learnt quite a bit software engineering knowledge from you. Georgeo, Wout, and Rita, thank you for providing invaluable suggestions.

I would like to thank all the ES group members and MOMA group members. Yaser, Onkar, Raghu, Susan, Guido, Rene, Menno, Boris, Rob, Nader, and Farid, discussions with you have been always helpful.

Thanks to you, Hendrik, Colin, Joost, Klaas, Michiel, Gert Jan, Mihai, Jo, Richard and Leonie. Thank you for all the discussions and suggestions that made me grow.

I also must thank you, Marion and Maria! Thank you for helping me out all the administrative matters and making my stay in the department and in the Netherlands a pleasant one. Thank you, Eva and Bowon, for your technique support on experiments.

My time in the Netherlands was made enjoyable in large part due to the many friends and groups that became a part of my life. We have so many groups for different activities. I am grateful for time spent with you. Georgeo, Boris, Menno, Onkar, Jani, Amit, Raghu, Jacob, Roderick, Matteo, Guido, Rene, Jan, Ila, I have had so much happy time with you during lunch, borrel, party, BBQ, gaming and so on. My time at Nijmegen was also enriched by the cuisine from different countries. Menno, Tatyana, Koen, Onkar, Jani, Amit, Abhishek, thank you for inviting me to taste the delicious food you made.

Outside Henk's lab, I had very fruitful collaborations with great scientists in Babraham Institute. Peter, Mikhail, and Steven, thank you for providing helpful advice and sharing your awesome pipeline.

It was also nice to have a Chinese community here. Jiankang and Qian, thank you for the food when I did not have a place to cook. Cheng, Ningqing, Tianran, Yuliang, Jieqiong, and Guoqiang, thank you for all the help. Every party with you makes me feel like in China. Ying, thank you for your support and love and thank you for taking care my life in the last year of my PhD. It is a treasure to meet you.

I also thank my old friends Wei, Zhong-Yin, and Yuan-Peng. We have been keeping in touch for so many years. Thank you for providing support and care all the way along.

Special thanks go to my family. Words cannot express how grateful I am to my mother and father. Thank you for all of the sacrifices that you have made on my behalf, and thank you for all of the unconditional love and care. I also thank my aunts, uncles and brothers for your love, care and support. I would not have made it this far without you. Grandpa, grandma, and dad, I miss you so much, and I really wished you could have been with me for my graduation. Thank you, my beloved dogs. I have not been able to see you for quite a long time, but every time when I see a dog somewhere the happy memories with you come rushing into my mind.

Thank all of you! If I have forgotten anyone, please accept my apologies.

I dedicate this thesis to everyone who loves me and everyone I love.

Curriculum Vitae

CURRICULUM VITAE

Shuang-Yin Wang was born on February 17th 1986, in Wuhan, China. He obtained his Bachelor degree in Biology and Master degree in Bioinformatics from Huazhong Agricultural University in 2009 and 2012, respectively. In September 2012, he started his PhD study in the laboratory of Prof. Henk G. Stunnenberg at Radboud Institute for Molecular Life Sciences, Randboud University Nijmegen, the Netherlands. His studies during PhD involved gene regulation at epigenetic and 3D genomic organization level. The results are described in this thesis.

Publications

PUBLICATIONS

Novakovic, B.*****, Habibi, E.*****, **Wang, S.-Y*.**, Arts, R.J.W., Davar, R., Megchelenbrink, W., Kim, B., Kuznetsova, T., Kox, M., Zwaag, J., et al. (2016). β-Glucan Reverses the Epigenetic State of LPS-Induced Immunological Tolerance. Cell *167*, 1354–1368.e14.

Joshi, O.*****, **Wang, S.-Y.***, Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell Stem Cell *17*, 748–757.

Kuznetsova, T.*****, **Wang, S.-Y.***, Rao, N.A.*****, Mandoli, A., Martens, J.H.A., Rother, N., Aartse, A., Groh, L., Janssen-Megens, E.M., Li, G., et al. (2015). Glucocorticoid receptor and nuclear factor kappa-b affect three-dimensional chromatin organization. Genome Biol. *16*, 264.

Liu, N.Q., Huurne, M. ter, Nguyen, L.N., Peng, T., **Wang, S.-Y.**, Studd, J.B., Joshi, O., Ongen, H., Bramsen, J.B., Yan, J., et al. (2017). The non-coding variant rs1800734 enhances DCLK3 expression through long-range interaction and promotes colorectal cancer progression. Nat. Commun. *8*, 14418.

Prange, K.H.M., Mandoli, A., Kuznetsova, T., **Wang, S.-Y.**, Sotoca, A.M., Marneth, A.E., van der Reijden, B.A., Stunnenberg, H.G., and Martens, J.H.A. (2017). MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. Oncogene.

Arts, R.J.W., Novakovic, B., ter Horst, R., Carvalho, A., Bekkering, S., Lachmandas, E., Rodrigues, F., Silvestre, R., Cheng, S.-C., **Wang, S.-Y.**, et al. (2016). Glutaminolysis and Fumarate Accumulation Integrate Immunometabolic and Epigenetic Programs in Trained Immunity. Cell Metab. *0*.

Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell *167*, 1415–1429.e19.

Chen, L., Ge, B., Casale, F.P., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yan, Y., Kundu, K., Ecker, S., et al. (2016). Genetic Drivers of Epigenetic and Transcriptional Variation in Human Immune Cells. Cell *167*, 1398–1414.e24.

Marks, H., Kerstens, H.H.D., Barakat, T.S., Splinter, E., Dirks, R.A.M., van Mierlo, G., Joshi, O., **Wang, S.-Y.**, Babak, T., Albers, C.A., et al. (2015). Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. Genome Biol. *16*, 149.

Wang, S.-Y., Cappellini, E., and Zhang, H.-Y. (2012). Why collagens best survived in fossils? Clues from amino acid thermal stability. Biochem. Biophys. Res. Commun. *422,* 5–7.
** Co-first authors.*